DOCUMENT RESUME

ED 329 452 SO 021 063

AUTHOR Holland, Simon
TITLE Artificial Intelligence, Education and Music: The Use of Artificial Intelligence To Encourage and Facilitate Music Composition by Novices.
INSTITUTION Open Univ., Walton, Bletchley, Bucks (England). Inst. of Educational Technology.
REPORT NO CITE-88
PUB DATE Jul 89
NOTE 305p.
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC13 Plus Postage.
DESCRIPTORS *Artificial Intelligence; Cognitive Processes; Cognitive Psychology; Computer Science; Computer Uses in Education; Educational Research; *Educational Technology; Educational Theories; *Musical Composition; *Music Education; Music Theory
IDENTIFIERS Harmony

ABSTRACT

The goal of this project was to find ways of using artificial intelligence to encourage and aid music composition by musical novices, particularly those without traditional musical skills. Two complementary approaches are presented. In the first approach, Harmony Space, a new kind of direct manipulation tool for music, is described. Two recent cognitive theories of harmony demonstrated how to design Harmony Space, with the expressivity to allow novices to sketch, analyze, modify, and compose harmonic sequences simply and clearly by moving two-dimensional patterns on a computer screen linked to a synthesizer. An investigation with a small number of subjects showed that Harmony Space considerably reduces the prerequisites required for novices to learn about, sketch, analyze, and experiment with harmony--activities that would normally be very difficult without considerable theoretical knowledge or instrumental skill. The second approach described is a knowledge-based tutoring system to help novices use an interface to compose chord sequences; this system is in the developmental stage. It is argued that traditional, remedial intelligent tutoring systems approaches were inadequate for tutoring in domains that require open-ended thinking. The foundation of a new approach is developed based on the exploration and transformation of case studies described in terms of chunks, styles, and plans. This approach draws on the characterization of creativity of Johnson-Laird (1988). Programs have been implemented to illustrate the feasibility of key parts of the new approach. (Author/DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED329452

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: SO
In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view. IR

Centre for Information Technology in Education

Artificial Intelligence, Education and Music

The use of Artificial Intelligence to encourage and facilitate music composition by novices

© Simon Holland

CITE Report No. 88

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY C. J. LAWLESS TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SO 021 063

Institute of Educational Technology
Open University
Walton Hall
Milton Keynes
Great Britain
MK7 6AA

BEST COPY AVAILABLE

Artificial Intelligence, Education and Music

The use of Artificial Intelligence to encourage and facilitate music composition by novices

Simon Holland

MSc Cognition, Computing and Psychology, University of Warwick, 1985
BSc Mathematics, University of Reading, 1976

Submitted for the degree of Doctor of Philosophy in Artificial Intelligence
The Open University, Milton Keynes

July 1989

© Simon Holland

Abstract

The goal of the research described in this thesis is to find ways of using artificial intelligence to encourage and facilitate music composition by musical novices, particularly those without traditional musical skills. Two complementary approaches are presented.

We show how two recent cognitive theories of harmony can be used to design a new kind of direct manipulation tool for music, known as "Harmony Space", with the expressivity to allow novices to sketch, analyse, modify and compose harmonic sequences simply and clearly by moving two-dimensional patterns on a computer screen linker to a synthesizer. Harmony Space provides novices with a way of describing and controlling harmonic structures and relationships using a single, principled, uniform spatial metaphor at various musical levels; note level, interval level, chord level, harmonic succession level and key level. A prototype interface has been implemented to demonstrate the coherence and feasibility of the design. An investigation with a small number of subjects demonstrates that Harmony Space considerably reduces the prerequisites required for novices to learn about, sketch, analyse and experiment with harmony - activities that would normally be very difficult for them without considerable theoretical knowledge or instrumental skill.

The second part of the thesis presents work towards a knowledge-based tutoring system to help novices using the interface to compose chord sequences. It is argued that traditional, remedial intelligent tutoring systems approaches are inadequate for tutoring in domains that require open-ended thinking. The foundation of a new approach is developed based on the exploration and transformation of case studies described in terms of chunks, styles and plans. This approach draws on a characterisation of creativity due to Johnson-Laird (1988). Programs have been implemented to illustrate the feasibility of key parts of the new approach.

Dedication

To Caroline, Simon, Peter, and my Mother and Father
with all my love

Acknowledgments

Thanks to my supervisor, Mark who has borne the brunt of the lunacy, the diversions, the modal harmonic ostinati, and kept me on planet earth throughout. Mark provided much needed support. He also provided an inspiring and energetic example of how to do outstanding research in Artificial Intelligence and Education. Without Mark's assured expertise in AI and his practical musicianship, I would not have had the confidence to start research in AI and Music.

This work would not have been started (or completed in a reasonable time scale) without Tim O'Shea, my co-supervisor. At the beginning, Tim found a simple way to get me to overcome my own doubts about tackling this challenging, interdisciplinary subject. At countless critical points, Tim solved practical difficulties and found just the right few words to motivate me. In the last few weeks before submission, Tim displayed exceptional generosity, horse sense, humour, intelligence and creativity. Tim invented ways of cutting the amount of work to do in half. He patiently and humorously picked garbage out of faxed drafts in the middle of the night. (The remaining garbage is due to my own cussedness). Tim is a first rate researcher and nurturer of research, but, more rarely, he is also a generous and open person.

I would like to thank Mark Steedman for his suggestion that Longuet-Higgins' theory might be a good area to explore for educational applications.

I would like to record my gratitude to George Kiss, a fine researcher and a conscientious teacher for being my mentor when I began Artificial Intelligence research.

Thanks to Claire O'Malley, who was supportive and generous in the last, crucial, phase of the write up, at a time when she had much more pressing demands on her time. I would like to thank people who gave their time to read drafts of the thesis, producing vital improvements; Peter Desain, Henkjan Honing and Sara Hennessey. Thanks to Fiona Spensley, Alistair Edwards, Mike Baker, John Sloboda, Trevor Bray and Mark Steedman for feedback on particular chapters or earlier drafts. Special thanks are due to Simon Bento for doing a wonderful job on proof-reading (any finally escaped typos are my fault - I look forward to reading Simon's thesis). Special thanks to Barry Jones for very kindly taking time to rescue me from various musical errors.

Thanks to Chris Fry, Kevin Jones and Tom Green who generously took time to give me useful pointers when I was getting in started in AI and Music. Thanks for encouragement, useful discussions, support and shared joy from other researchers in AI and music and other (more or less) related fields; Dave Levitt, Henry Lieberman, folk at the Media Lab; Peter Desain, Henkjan Honing, Sterling Beckwith, David Wessel, Lelio Camilleri, Linda Sorisio, Christopher Longuet-Higgins, Jeanne Bamberger, Wally Feurzeig, Gerald Balzano, Mike Baker, Eric Clarke, Ian Cross, Stephen Pope, Mira Balaban, Kemal Ebcioglu, Rick Ashley,

Derek Ludwig (and others at NorthWestern); Colin Wells, Nigel Morgan, Keith Waters, Anders Friberg, Lennart Fahlen, Steve McAdams, Xavier Rodet, Jean Baptiste Barrière and others at IRCAM; Marilyn Taft Thomas, Roger Dannenberg and folk at Carnegie Mellon; Stellan Ohlsson, Jeff Bonar and others at LRDC; Alan Borning, Bill Gaver, Fred Lerdahl, Alan Marsden, Stephen Page, Mike Greenhough, Christoph Lischka, Neil Todd, Tony Watkins, Simon Emmerson and Pete Howell.

Thanks for strong practical and moral support to Eileen Scanlon, Diana Laurillard, Pete Whalley, Sterling Beckwith, Nancy Marlett, Rachel Hewson, Richard Southall, and Jayalaksmi.

Special thanks to Olwyn Wilson, Dave Wakely (watch out, National Sound Archive), Tony Durham, Conrad Cork, Phil Butcher, Dianne Murray, Steve Cook, Mike Brayshaw, Di Mason, Pat Cross and Rick Evertsz.

Thanks for various kinds of help and encouragement to Pat Fung, Fiona Spensley, Anne Blandford, Norman White, Gordon Burt, Nicola Durbridge, Kathy Evans, Rae Sibbitt, Roshni Devi, Laurence Alpay, Rod Moyse, Betty Swift, Kate Stainton-Ellis, Beryl Crooks, Anne Wood, Margaret Harkin, Helen Boyce, Arthur Stutt, John Domingue, Tony Hasemer, Tim Rajan, Marc Eisenstadt, Hank Kahney, Don Clarke, Yibing Lee, Ann Jones, John Close, Richard Joiner, Bradley, Ed Lisle, Mike Fox Jitu Patel, Dave Kay, Simon Nuttall, Phil Swann, Benedict Heal, Hansa Solanki, and *all* in the Centre for Information Technology in Education, the Institute for Educational Technology, the Computer-Aided Learning Research Group and the Human Cognition Research Laboratory.

Thanks to various musicians who helped: Herbert Balsam, Martin Brems, Tony Jefferies, Jonathan Brown, Grant Shipcott, Sean Graham, John Winter, Guillame Ormond, John Walsh, John Close, Mike Baker, Rod Moyse, John Bickersteth, Fred de Borst, Gerda Koetje, Dave Gollancz, Richard Middleton, Donald Burrows, Rev. Cousins, Dean Lloyd, Frank Hawkins and Chris Clark.

I owe a considerable typographical debt to Rob Waller and Rachel Hewson.

Emergency squad: very big thanks to Simon Bento, Pete Whalley (not for the first time), Claire O'Malley, Katerine Bielaczyc, Tom Vincent and Aileen Treacy.

Utterly final emergency squad: Tim O'Shea, Claire O'Malley, Simon Bento, Sara Hennessy, Rae Sibbitt, Roshni Devi, Benedict Nugroho Budi Priyanto, Dana Sensuse, Bill Ball, Claudine and the security team. A special thank you to Dave Perry for help at this time.

Thanks to the experimental subjects for their hard work.

Thanks to Peter Mennim, Anne, Gerry, Rachel, Christine, Lindsay, Noleen, Sinead, Aine, Steven Brown, Richard Wrigley and Penelope for injecting a little sanity.

Bill Harpe, Wendy Harpe, Stevie Smith, Martin Brems, Judy Bates, Sally Morris, Niel, Stephen Knox, Eddie Tagoe, Rie Toft Hansen and Vivi Mitts for alternative reality training and companionship.

Typhun, Beyhan and Dr. Torsun for initial encouragement.

Thanks and apologies to those who accidentally aren't thanked properly by name.

Big thanks to Lydia, Tina, Howard, Ralph, Claire, Phillipa, Andrew, Thomas, Jo, Ursula, Christine, Barbara, Sharon, Phillip (first novice Harmony Space user), David, Francis, Julie and all of the family.

The most heartfelt thanks of all are due to my wife Caroline. This research would not have been possible at all without her sustained love, support, sympathy, humour and help. Lots of love.

Thanks to my sons Simon and Peter for love, help, fun, education and for reminding me that there are things infinitely more important than research. Big hugs.

Thanks to my mother and the memory of my father. Both are fine musicians and exceptional people whom I love, admire and respect deeply. Lots of love.

This research was supported by ESRC studentship C00428525025. The OU research committee gave an equipment grant. IET gave a study extension. These are all gratefully acknowledged.

Contents

PART 1 INTRODUCTION

Chapter 1 Introduction

1 The problem
2 General approach
3 Relevance for ITS in domains other than music
4 Timeliness
5 Roots of the research
6 Structure of the thesis

Chapter 2 Early uses of computers in teaching music

1 Computer-aided instruction in music education
2 Music Logos
3 Tools for the student
4 Interactive graphic music games
5 Conclusions on early approaches

Chapter 3 The musician-machine interface

1 The importance of the interface
2 General strategies for designing good interfaces for novices
3 MMI meets HCI - a two way traffic
4 Intelligent Instruments
5 MMI in music education
6 Conclusions on the musician-machine interface

Chapter 4 Intelligent Tutoring systems for music composition

1 A role for the 'traditional' ITS model in music composition
2 Vivace: an expert system for harmonisation
3 MacVoice: a critic for voice-leading
4 Lasso: an intelligent tutoring system for 16th century counterpoint
5 Conclusions on intelligent tutors for music composition
6 Related work in other fields
7 Conclusions on the use of computers in teaching music

PART II HARMONY SPACE

Chapter 5 Harmony Space

1 Longuet-Higgins' theory of harmony
2 The 'statics' of harmony
3 A computer-based learning environment
4 Representing harmonic succession
5 Analysing music formally in Harmony Space
6 Conclusions

Chapter 6 Adapting Longuet-Higgins' theory for Harmony Space

1 Longuet-Higgins' non-repeating space
2 Comparison of the two Longuet-Higgins' spaces
3 Direct manipulation
4 Use of the 12-fold representation in Harmony Space
5 Conclusions

Chapter 7 Harmony Space using Balzano's Theory

1 Balzano's theory of Harmony
2 Review of design decisions
3 Harmony Space using Balzano's theory
4 Expressivity of a version of Harmony Space based on thirds space
5 Analysing a chord progression in the thirds space
6 Relationship between the alternative Harmony Space designs
7 Implementation of Harmony Space
8 Some educational uses of Harmony Space
9 Intended users of the tool
10 Related interfaces
11 Limitations and further work
12 Conclusions

PART III TOWARDS AN INTELLIGENT TUTOR FOR MUSIC COMPOSITION

Chapter 8 An architecture for a knowledge-based tutor for music composition

1 Introduction
2 Tutoring in open-ended creative domains
3 Architecture of a knowledge-based tutor for music composition
3.1 An interaction style: the 'Paul Simon' method
3.2 Representation of pieces
3.3 Representation of styles
3.4 Musical plans
4 ITS perspectives on the framework
5 Partial implementation of MC
6 Further work: supporting active teaching strategies.
7 Problems and limitations.
8 Summary and conclusions

PART IV CONCLUSIONS

Chapter 13 Novices experiences with Harmony space

1 Introduction
2 Method
3 Session 1: YB: harmonic analysis by mechanical methods
4 Session 2: BN: an eight year old's experience with Harmony Space
5 Session 3: RJ: virtuoso visual performance
6 Session 4: HS: analysing Mozart's Ave Verum
7 Session 5: LA: putting it together and making sense of the sounds
8 Conclusions
9 Extensions
10 Further work

Chapter 14 Conclusions and further work

1 Contribution
2 Criticisms and limitations
3 Further work

References

Appendices

1 Harmony Space Prototype implementation
2 Harmony Space notional interface
3 Chord symbol conventions

TABLE OF FIGURES

Chapter 2

Fig 1 Guido intervals program
Fig 2 Guido melody program
Fig 3 Guido chord quality program
Fig 4 Guido harmony program
Fig 5 Guido rhythm program
Fig 6 Fast regular beat
Fig 7 Slow regular beat
Fig 8 "London Bridge is falling down"
Fig 9 "London Bridge is falling down" (with pulse)
Fig 10 "London Bridge is falling down" (with pulse & linked trace)
Fig 11 Lamb's interactive graphics game: initial set up
Fig 12 Player draws a curve
Fig 13 Curve is straightened to preserve left-to-right flow
Fig 14 Computer changes curve into discrete notes
Fig 15 Circled notes are to be transformed
Fig 16 Notes to be transformed are displayed as hollow triangles

Chapter 3

Fig 17 Musician/score/instrument relationship 1
Fig 18 Musician/score/instrument relationship 2
Fig 19 Musician/score/instrument relationship 3
Fig 20 Blank working screen for Hookup
Fig 21 Hookup with various icons waiting to be hooked up
Fig 22 A Hookup program to play chord progressions
Fig 23 A typical 'M' screen
Fig 24 Jam Factory

Chapter 4

Fig 25 MacVoice

Chapter 5

Fig 1 Longuet-Higgins' note array
Fig 2 12-fold version of Longuet-Higgins note array
Fig 3 Triads in 12-fold space
Fig 4 Scale-tone sevenths in 12-fold space
Fig 6 Cycles of fifths in 12-fold space
Fig 7 A scalic progression
Fig 8 A chromatic progression
Fig 9 Diatonic cycles of thirds
Fig 10 Movement in major thirds unconstrained by key
Fig 11 Movement in minor thirds unconstrained by key
Fig 12 Table of intervals in 12-fold space
Fig 13 Chord progression of "All the Things You Are" in the 12-fold space

Chapter 6

Fig 1 Longuet-Higgins original non-repeating space
Fig 2 VI triad in the non-repeating space
Fig 3 II triad in the non-repeating space
Fig 4 Triads in C major in the non-repeating space
Fig 5 Comparison of chromatic progressions in the non-repeating and 12-fold spaces

Chapter 7

Fig 1 The semitone space
Fig 2 The fifths space
Fig 3 Steps of tones and tritones
Fig 4 Steps of major thirds and minor thirds
Fig 5 Diatonic scale embedded in fifths space
Fig 6 Balzano's thirds space
Fig 7 Harmony Space using thirds representation
Fig 8 Triads in the thirds space
Fig 9 Scale tone sevenths in the thirds space
Fig 10 Trajectories down the fifths axis in the thirds space
Fig 11 Scalic trajectories
Fig 12 Chromatic trajectories
Fig 13 Trace of "All The Things You Are" in the thirds space
Fig 14 Screen dump from Harmony Space prototype
Fig 15 Longuet-Higgins' Light organ
Fig 16 Harmony Grid

Fig 17 Music Mouse

Chapter 8

Fig 1 General architecture of proposed tutoring system
Fig 2 Kinds of domain knowledge in MC architecture
Fig 3 Chord sequence of Randy Crawford's (1978) "Street Life"
Fig 4 A version of "Abracadabra" (The Steve Miller band, 1984)
Fig 5 "Abracadabra" under assembly - motor rhythm stage
Fig 6 "Abracadabra" reconstructed in seven chunks
Fig 6a Major version of "Abracadabra"
Fig 6b Version of "Abracadabra" in ninths
Fig 6c Version of "Abracadabra" with chord alphabet restriction lifted
Fig 6d Version of "Abracadabra" with new harmonic trajectory direction
Fig 7 A prototypical reggae accompaniment pattern
Fig 8 A prototypical Fats Domino pattern
Fig 9 Tree of songs instantiating the "return home conspicuously" plan
Fig 10 First eight bars of "All the things you are"
Fig 11 Example modal replanning of "All the things you are"

Chapter 9

Fig 1 Harmony Space with chord function labelling
Fig 2 Harmony Space with semitone number labelling
Fig 3 Harmony Space with alphabet name labelling

Chapter 10

Fig 1 Suggested Harmony Space metric display
Fig 2 Assessing the harmonic resources of the pentatonic scale

14

Chapter 1: Introduction

The goal of this research is to find ways of using artificial intelligence to encourage and facilitate the composition of music by novices, particularly those without traditional skills. It will be argued that the research also has implications for intelligent tutoring systems in general, because the fostering of open-ended approaches is relevant to any sufficiently complex domain (e.g. engineering design, architectural design, etcetera).

1 The problem

The goal of music composition can vary from composer to composer and occasion to occasion and cannot normally be stated precisely beforehand. Almost the only goal that applies to music composition in general is "compose something interesting" (Levitt, 1985). The original focus of the research was on novices who have little or no formal education in music, and are teaching themselves, perhaps with the help of their friends, outside a formal educational setting. As Camilleri (1987) points out in a critical review of an earlier report on part of this research (Holland,1987a), this aim fits well with the general educational orientation of the institution which has supported the research (the Open University). Following this original aim, our illustrations and vocabulary are drawn predominantly from popular music and jazz, since it seems good educational practice to focus on music that is highly motivating to potential students. In fact, it turns out that most of the work is equally applicable to western tonal music in general.

2 General Approach

2.1 Introduction to the approach

Much of the current work in open-ended computer-based learning environments still

draws its inspiration ultimately from work on the Logo project (Papert, 1980). Logo is a programming language that attempts to link, amongst other things, the idea of a function to a child's intuitions about her own body movement. The validity of the claims made for the educational effectiveness of Logo are now hotly contested (see, for example, Kurland, Pea, Clement and Mawby (1986)). However, it can at least be claimed with some justice that Logo bears a special relationship to mathematics due to the more or less central role played by the mathematical function in both mathematics (Kasner and Newman, 1968) and Logo (Hayes and Sutherland, 1989). In many other domains, such as music, it is far less clear that functional application is a particularly appropriate metaphor. When designing an exploratory microworld for some new domain, we advocate that metaphors or formalisms should be sought that lie at the heart of *that particular domain* in order to be given appropriate computational embodiment. This advocated approach, which we will illustrate in Part II contrasts with the more 'traditional' microworld approach of addressing a non-mathematical domain by adding superficial, domain-specific facilities to a Lisp-like programming language.

2.2 Developing a learning environment appealing to intuitions in several modalities

Another source of power in microworlds can arise from linking a domain metaphor to simple intuitions in many modalities (sound, gesture, animation etc.). The learning environment for harmony discussed in part II of the thesis links cognitive theories of harmony at a number of levels to intuitions in three modalities; visual, auditory and body movement. This learning environment uses the cognitive theories to make abstract harmonic concepts concrete and manipulable. This allows abstract concepts to be treated in a computer -based environment using all the advantages of direct manipulation methods (Shneiderman, 1982), (O'Malley, in press).

2.3 The need for new intelligent tutoring systems approaches in domains that demand creative behaviour

One of the most interesting consequence of the choice of domain from an intelligent tutoring systems (often abbreviated to 'ITS') point of view is that it is inappropriate for the system to be directly prescriptive or critical. There already exist a number of

rule-based intelligent tutors for music composition which have a prescriptive and critical style. For example, Stephen Newcomb's (1985) LASSO is a tutor for a paradigm or style of music known as 16th Century Counterpoint, and Marilyn Taft Thomas's (1985) MacVoice is an intelligent tutor for a musical task known as voice leading - which is part of the related task of four-part harmonisation. Both of these pioneering musical Intelligent Tutoring Systems are prescriptive and critical and it is appropriate for their particular domains that they should use this approach. Why then could not a tutor for popular music composition be prescriptive in the same way? The answer is that 16th Century Counterpoint and four part harmonisation are genres that (as taught in schools) can be considered to have stable and known rules, goals, constraints and bugs.[1] In the case of popular music it might be possible to write a series of distinct expert components for a range of genres or paradigms of composition but this would not satisfactorily address the problem. This is particularly so in the case of self educators driven by personal motivation rather than a need to satisfy an examination or course requirement. The reasons for this are twofold. Firstly, genres in popular culture are still rapidly evolving from a rich set of background genres. A new song can easily combine features, rules and goals from genres not previously mixed. Designing a series of expert components that could perspicaciously characterise each genre or paradigm, be freely interchangeable and clearly exhibit inter-relationships between the styles, remains an unresolved research problem. Secondly, any new piece of music worth its salt invents new rules, goals and constraints specifically for that piece, and may consequently violate some rules normally expected of the genre. In this constantly evolving domain it is hard to see how any intelligent tutoring system could know enough to be prescriptive or critical without the risk of misunderstanding all the most innovative and interesting ideas.

If we are claiming that a prescriptive or critical style is not appropriate for the domain, does this mean that we are limited to a non-directed microworld or environment approach? To the contrary, we will argue that microworlds are useful for such tasks, but that they need complementing with other ITS approaches.

[1] This is an over-simplification because as we will note in chapter 4, work on formal models of musical processes has shown that text book rules are grossly inadequate - but the assumption of a stable, understood style is often made by educators.

2.4 A framework for individually relevant guidance

An important issue in the design of microworlds is the issue of providing individually relevant guidance. For example, the music microworld Music Logo (Bamberger, 1986), comes with a well-written manual containing graded work sheets. Anyone, novice or expert who worked through the book would learn a lot about music or see it anew from a fresh angle - but the issue of educationally exploiting individual differences in areas of motivation and ability is not considered. In a school context, this is easily addressed - the teacher can consider each individual's special interests and point them to different worksheets in different orders. She can suggest modified versions of the exercises to take into account likes and dislikes - so maximising motivation. In the case of a self-educator, opportunities are missed to greatly accelerate the learning process and fuel motivation.

2.5 Broad architecture of a tutor for music composition

Given that we cannot be overtly prescriptive in our domain but want to open up the possibility of tackling the issue of providing individually relevant guidance, how can we proceed, since neither the open-ended microworld approach nor the traditional interventionist tutoring approach is adequate for both purposes? In broad terms, our answer is to use a non-prescriptive knowledge-based tutor linked to one or more musical microworlds. The point of such an architecture is that the knowledge base could provide abstract knowledge and individually relevant guidance for exploring the microworlds, and the microworlds give principled visual and gestural alternative representations to auditory experience and predominantly textual, conceptual knowledge. This kind of approach coincides with the broad framework of Guided Discovery Learning, proposed by Elsom-Cook (1984) as a framework for combining the benefits of learning environments with those of interventionist tutors. Broadly speaking, Guided Discovery Learning methodology proposes the framework of providing a microworld in which discovery learning can take place, integrated with a tutor that is capable of executing teaching strategies that range from an almost completely non-interventionist approach to highly regimented strategies that give the tutor a much larger degree of control than the student. The place of the

ITS architecture of this project within the Guided Discovery Learning paradigm is discussed in Holland and Elsom-Cook (in press). Of course, Guided Discovery Learning does not in itself solve the problem of tutoring in open-ended domains as outlined earlier, but it does give a framework for combining discovery learning in a computer environment with the provision of individually relevant guidance.

Foundations for a new tutoring approach for open-ended domains are developed based on the exploration and transformation of case studies described in terms of chunks, styles and plans. This approach draws on a characterisation of creativity due to Johnson-Laird (1988). Interventionist teaching strategies for providing individually relevant guidance are not presented in the thesis. The new tutoring framework looks very well suited to supporting a wide range of active teaching strategies but the work is concentrated on finding solutions for the core problems that had to be solved to make a tutor possible at all: the devising of ways of characterising strategic considerations about chord sequences in ways that are both suitable to the modelling of creative processes and at the same time potentially communicable to novices. Although a fully implemented tutoring system is not described, programs to illustrate the feasibility of key parts of the new approach have been implemented. One of the main techniques developed is to describe chord sequences in terms of plans and manipulate the plans using a musical planner, which we will now discuss.

2.6 Complementing microworld representations with abstract knowledge

A major issue related to the issue of individually relevant guidance (and one that microworld approaches in general have only recently started to tackle) is the issue of providing ***abstract knowledge*** to complement direct experience and to guide exploration of the domain. The need for this can be illustrated in the case of music composition as follows. Music composition is about more than just familiarity with raw musical materials and the way they are perceived. When we listen to any piece of music, we are influenced by expectations based on other pieces of music, and by styles that we already know. To explore effectively ways of using the raw materials, we need some way of relating them to musical knowledge about existing works, styles and sensible musical behaviour in general. We introduce a new way

of characterising abstract knowledge about harmonic sequences called "musical plans", to try to make explicit plausible reasons why existing pieces use raw materials in the way they do, and to characterise a range of common intelligent musical behaviors. Musical plans suggest various ways of organising raw musical materials to satisfy various common musical goals.

3 Relevance for ITS in domains other than music

Almost all work in intelligent tutoring systems up until now has been restricted to domains such as arithmetic, electronics, physics and programming in which there are clear-cut goals and in which it is relatively easy to test formally if a solution is right or not. In domains demanding creative behaviour, on the other hand, goals may not be precisely definable before the event, and "good" solutions may be very hard to identify formally. We will argue that existing ITS approaches are inadequate for such domains, and that finding ITS methodologies for dealing with open-ended domains of this nature is an important task for ITS as a whole. Tutoring creative behaviour cannot be dismissed as being relevant to only a handful of "unusual" domains; to the contrary, this issue is of central concern if intelligent tutoring systems are ever to progress beyond highly restricted domains into real world areas with less well-defined goals and methods. We can postulate that in any knowledge-rich domain with a sufficiently large search space, searching the space can be treated as an open-ended creative task. Any domain at all, even arithmetic, electronics, physics or programming may need to be considered creatively when a requirement arises to extend the domain or reconceptualise it. Indeed, some teachers and psychologists might argue that extension and reconceptualisation lie at the heart of effective learning in general. For these reasons, progress towards architectures for tutoring creative tasks has implications for the discipline as a whole. This project does not solve the problem of tutoring in a creative domain in general, but makes some first steps.

A second reason why intelligent tutoring systems for music composition may have important implications for the discipline as a whole, is that due to the demanding nature of the domain, research in ITS for music can act as a valuable *forcing function* for ITS methodology and architectures, introducing unusual perspectives

and fresh techniques.

4 Timeliness

The research is timely in a practical sense due to the coincidence of a number of theoretical, educational and technological developments. AI researchers have taken an active interest in AI and music research for some time; examples include Simon (Simon and Sumner, 1968), Winograd (1968), Minsky (1981a), Longuet-Higgins (1962), Steedman (1984) and Johnson-Laird (1988). This work has been very broad, ranging from surprising strides in clarifying music theory (Longuet-Higgins, 1962) and ways of relating music to other areas of human cognition (Lerdahl and Jackendoff, 1983), to finding flexible ways of representing general time-embedded highly interacting parallel processes (Rodet and Cointe, 1984). More recently, there has been a strong upsurge of research in the psychology of music, particularly research influenced by the cognitive paradigm (Sloboda, 1985). This has been associated with a shift of theoretical attention from low level phenomena to higher level musical processes in rich contexts. This has provided a stream of powerful explanations and theories for musical phenomena that had previously defied *formal* analysis for centuries, e.g. aspects of ;

- harmony (Longuet-Higgins, 1962), (Balzano, 1980),
- rhythm (Longuet-Higgins and Lee, 1984),
- voice-leading (Wright, 1986), (McAdams and Bregman, 1985), (Cross, Howell and West, 1988), (Huron, 1988)
- hierarchical reduction (Lerdahl and Jackendoff, 1983),
- rubato (Todd, 1985).

Up until recently, much of this research had very little scope for application due to the rarity and expense of computer controllable musical instruments. The recent emergence of inexpensive high quality electronic musical instruments in the marketplace, coinciding with the recent establishment (Loy, 1985) of an industry standard (MIDI) for connecting computers to electronic musical instruments means that all of this work is now open for the first time to widespread practical

application. These converging considerations mean that work done now in ITS and music is timely and capable of exploitation.

5 Roots of the research

Part II of the thesis was inspired directly by a single piece of work - Longuet-Higgins (1962). Some points of the interface design were influenced by Smith, Irby, Kimball, Verplank and Harslem (1982), and a later version of the interface was designed to take into account Balzano (1980). Part III of the thesis (the architecture for a tutoring system for music composition) was loosely inspired by work in machine learning due to Lenat (1982), but arose directly from consideration of the 'traditional' structure of intelligent tutoring systems, as described in, for example, Sleeman and Brown (1982), and simple observations about the nature of open-ended tasks. Inspiration for the constraint-based planner for music composition came from Levitt's work (1981, 1985).

6 Structure of thesis

The thesis is divided into 4 sections, together with some appendices.

Part I provides an introduction to the thesis as a whole. This part includes a selected review of previous work related to Artificial Intelligence, Education and Music.

Part II reviews two cognitive theories of harmony and shows how they can be used to design learning environments for tonal harmony. Some programs developed independently that relate to the work are discussed.

Part III outlines an architecture for tutoring music composition that combines the learning environment with sources of more abstract knowledge. A musical planner is the key abstract knowledge component of the architecture.

Part IV outlines a practical investigation carried out with Harmony Space, summarises the results of the thesis as a whole and looks at possibilities for further work.

Chapter 2: Early uses of computers in teaching music

In Part I of the thesis, we review the uses of computers in music education. Chapter 2 looks at early uses of computers in music education. Chapter 3 reviews recent trends in the musician machine interface. Chapter 4 is concerned with intelligent tutors for music. The final chapter includes pointers to related work in other fields.

In this chapter, we review early uses of computers in music education, with special attention to the needs of novices learning to compose. We review four representative and contrasted early approaches. The four approaches considered are: computer-assisted instruction (CAI) in music; music logos; tools for music education; and interactive graphical games. (Straightforward music editors and computer musical instruments are omitted, except for brief references.) The review is selective throughout. Applications may display characteristics of more than one division, so the categories cannot be taken as hard and fast. The structure of the chapter can be summed up as follows;

1 Computer-aided instruction in music education
2 Music logos
3 Tools for the student
4 Games
5 Conclusions on early approaches

We will begin by considering the first of the four early approaches, namely computer-aided instruction in music education.

1 Computer-aided instruction in music education

Historically, the first use of computers in teaching music or teaching any other

subject for that matter, was associated with programmed learning, derived originally from behaviourism. Branching programs, perhaps the most common type of programmed learning, step through essentially the following algorithm. (The next two or three paragraphs draw on the best existing survey of the area, to be found in O'Shea and Self (1983)).

1 Present a "frame" to the student i.e.
   Present the student with prestored material (textual or audio visual),
   Solicit a response from the student.
2 Compare the response literally with prestored alternative responses.
3 Give any prestored comment associated with the response.
4 Look up the next frame to present on the basis of the response.

Branching programs can be used to present a kind of individualised tutorial in which different answers would cause students to be treated differently. However, every possible path through the frames must be considered by the author of the material, or else the material might not make sense following some routes. The number of possible routes can quickly become very large, making this kind of material very laborious to prepare (O'Shea and Self, 1983).

A development of programmed learning, called 'generative computer assisted learning', allowed questions and answers to be generated from a template instead of being pre-stored. For example, new equations to be solved could be generated and solved (for testing the response) or new sequences of chords to be identified could be generated by a grammar (Gross, 1984). The student's results could be used to control the difficulty and subject of the next example according to some simple prespecified strategy. Generative computer assisted learning was often used for drills rather than tutorial presentations.

Music educators such as Gross (1984) and Hofstetter (1975) used systems along these lines to attempt to teach music fundamentals, music theory, music history and aural training. Hofstetter's GUIDO system at the University of Delaware (Hofstetter 1975 and 1981), is a paradigmatic example of computer aided instruction in ear

training.

1.1 GUIDO:- Ear training by computer administered dictation

The GUIDO system (named after the 11th Century music educator Guido d'Arezzo) has a twin purpose; the provision of computer based ear training, and the collection of data for research into the acquisition of aural skills. GUIDO can oversee dictation in five areas; intervals, melody, chords, harmony and rhythm. GUIDO plays musical examples by means of a four voice synthesizer, and the student is invited to select from a multiple-choice touch-sensitive display what he thinks he has heard. Example displays from each of the five dictation areas are reproduced in Figs 1 - 5.

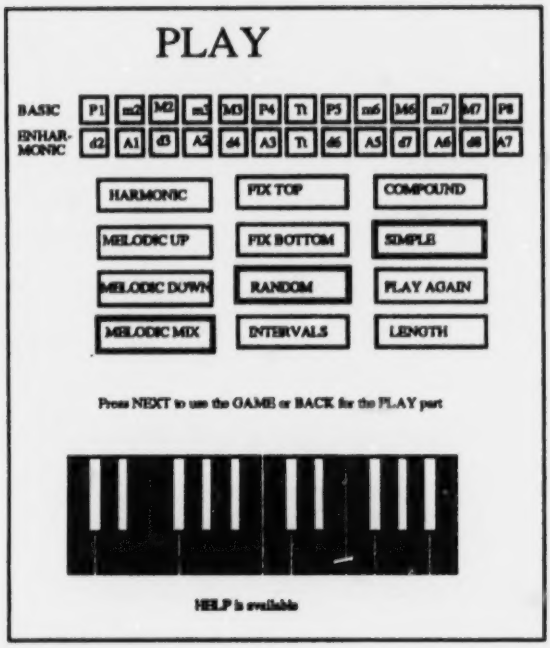


Fig 1 Guido intervals program

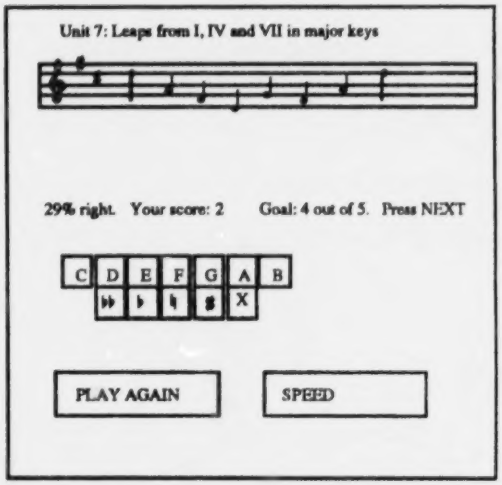


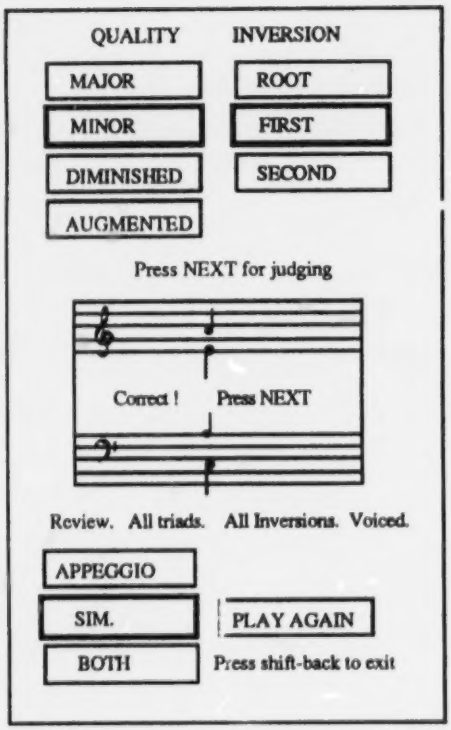
Fig 2 Guido melody program

Both figures redrawn from Hofstetter (1981) pages 85 and 86.

Originals © University of Delaware Plato Project.

We will look at each of the drills in a little detail. (The following descriptions of the various drills draw closely on Hofstetter (1981)). For the interval drill (fig 1), GUIDO plays an interval, and the student touches a box on the screen to indicate what interval was heard. The three columns of boxes in the middle of the screen can

be used by the teacher or student to control how the dictation is given. The teacher has the option of "locking" the control boxes or allowing the student to control them.



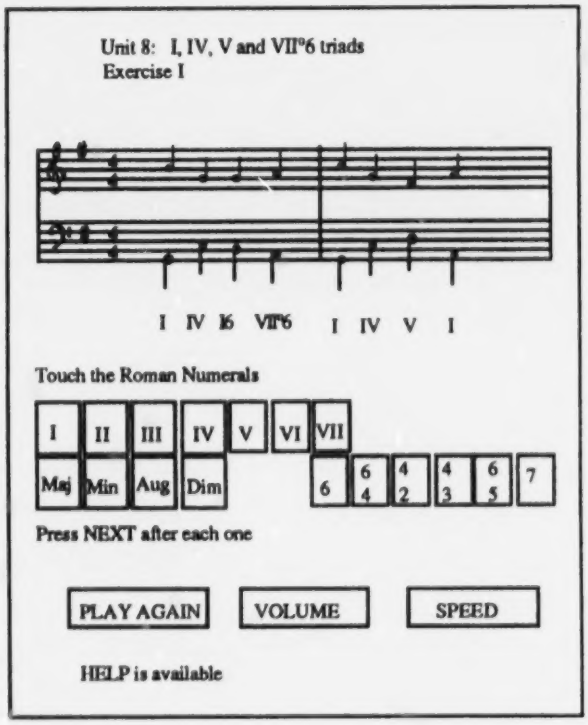


Fig 3 Guido chord quality program Fig 4 Guido harmony program

Both figures redrawn from Hofstetter (1981) page 86.

Originals © University of Delaware Plato Project.

The legends (fig 1) are mostly self explanatory, and all details do not need to be grasped for our purposes, but two non-obvious points are worth mentioning. Firstly, the box marked "intervals" can be used to eliminate from the dictation any selected intervals; secondly, the attractive looking keyboard display at the bottom of the screen is used in this program only when fixing the lower or upper note of intervals. In the melody program (fig 2), GUIDO plays a tune and the student tries to reproduce it by touching the note "boxes". Options are available for the student to input using boxes marked with solfa syllables, or the keyboard shown in fig 1. The students' input is sounded and displayed simultaneously as staff notation.

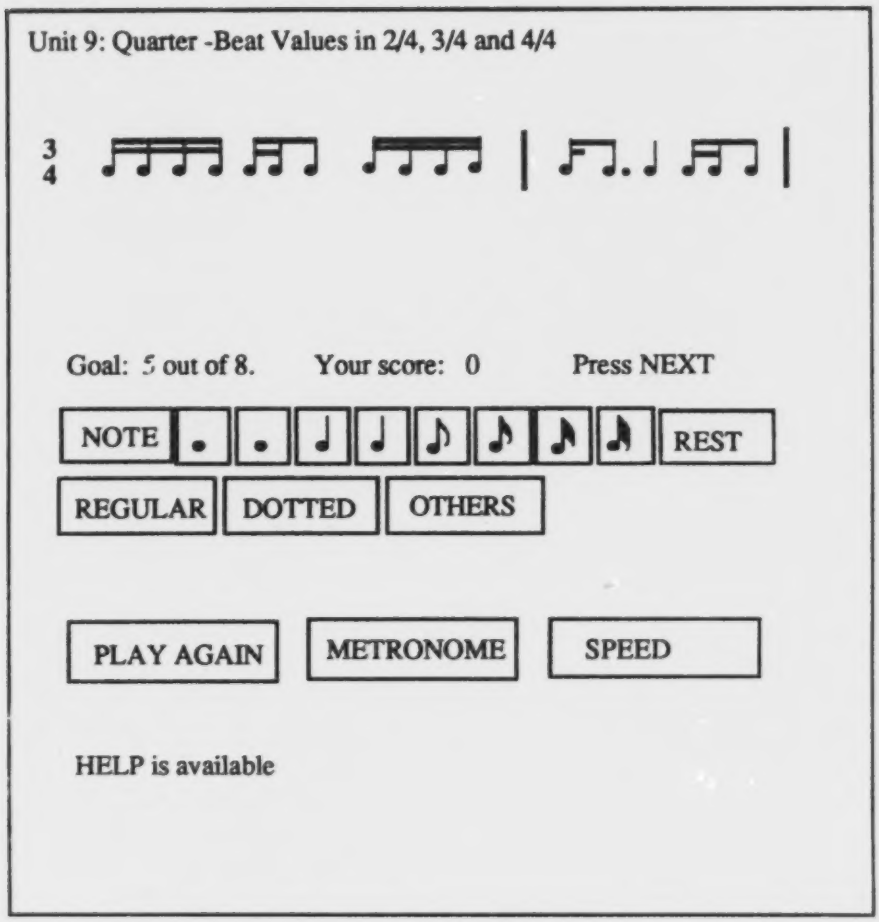


Fig 5 Guido rhythm program

Redrawn from Hofstetter (1981) page 87.

Original © University of Delaware Plato Project.

The chord quality program asks the student to indicate the quality and inversion of chords sounded (fig 3). In the harmony program (fig 4), the student listens to a four part exercise in chorale style, and then indicates what he has heard by touching boxes for Roman numerals and soprano and bass notes. GUIDO displays the responses in staff notation. As in most of the programs, the student can ask to hear the example played again, can control the tempo, and can alter the relative loudness of the different voices. Finally, in the rhythm dictation program (fig. 5), GUIDO plays a rhythm, and the student touches boxes to indicate what was heard. GUIDO displays the response as it is input.

1.1.1 Discovery learning mode

GUIDO does not adhere strictly to programmed learning dogma. All of the GUIDO programs can be used for discovery learning as well as dictation. In discovery learning mode, the touch sensitive displays can be used playfully to "play" GUIDO as a musical instrument.

1.2 Advantages of Computer Aided Instruction in music dictation/aural training

If the aim is to practise and test aural skills formally, then dictation, whether administered by teacher or computer is a useful method. GUIDO's competitors for dictation in undergraduate courses are classroom dictation and tape. The advantages claimed for GUIDO are firstly that speed of dictation, order of presentation and time allowed for answering can be individualised for each student; secondly that the students replies can be recorded, analysed and acted upon instantly. Hofstetter demonstrated experimentally that a group of students using the GUIDO harmony program scored better in subsequent traditional classroom dictation using a piano than a control group using taped dictation practice (Hofstetter, 1975, page 104). The vast majority of students seem to enjoy using GUIDO (Hofstetter 1975).

The criticisms of traditional CAI in fields *other* than music can be summarised as follows. The system is neither able to recognise any correct answers that were not catered for explicitly by the author, nor to make educational capital out of slightly wrong answers, or answers that are "wrong for a good reason" that were not forseen. If a student happens to know most of the material in the course, and simply wishes to extract one or two pieces of information, many programmed learning systems cannot help; they can only run through those paths the author allowed for. CAI in the restricted domain of aural dictation seems to avoid most of these charges. Many aural drills only have one right answer, and GUIDO often allows a choice of ways of presenting this answer. It is not always obvious how educational capital could be made out of wrong answers to aural dictation. Certainly, with all the control boxes off, GUIDO gives the student full rein to extract "information". Aural training seems to be one of the few areas where traditional CAI really is an

improvement over the alternatives. (It might be argued that there are better ways of helping someone to learn how to recognise, say, chord qualities, than by dictation, but that is a different question).

1.3 Limitations of Computer Aided Instruction in music

For our purposes, 'traditional' CAI in music to date has at least two limitations. Firstly, in practice most CAI in music education has been strictly limited to music fundamentals (i.e simple questions and answers about basic concepts), music history or aural training. Gross (1984), using an aural dictation system similar to GUIDO found that while practice in aural skills with such systems made the students better at low level aural skills , the effect on larger scale topics was not measurable. Hence a CAI program for aural dictation might make students better at taking aural dictation, but would be likely to have minimal impact on composition skills. Secondly, the problems of trying to design a 'traditional' CAI program to teach composition skills would be hamstrung by the complex open ended nature of the task. It seems unlikely that the mere presentation of text and audio visual material would go very far towards teaching anyone how to compose in an open-ended sense. CAI depends on being able to anticipate all possible responses, whereas one essential component of composition can be the unexpected.

2 Music Logos

Music Logo, as a generic approach, contains elements of music editors and games. However, its intellectual parentage and ambitious aims are such that the approach can reasonably be discussed in a category of its own. In its original and pioneering incarnation, Music Logo (Bamberger, 1974) took the form of research revolving around prototype versions of the "Tune Blocks" and the "Time Machine" games, which we will describe shortly. More recently, Music Logo has become commercially available as a proprietary programming language (Bamberger, 1986), running on Apple II computers. In the suggested pedagogical uses of the commercial computer language, versions of "Tune Blocks" and the "Time Machine" play central roles. But Music Logo as a generic approach is an attempt to apply the ideas of Logo to music, so it is fruitful to begin by asking "What is Logo?".

2.1 What is Logo?

Logo is a programming language that tries to embody an educational philosophy of "learning without being taught" (Papert, 1980, page 7). Logo is an attempt to link deep and powerful mathematical ideas (e.g. the idea of a function) to a child's intuitions about her own body movement (Papert, 1980, page viii). Logo draws on many different communities and bodies of work: Feurzeig, Piaget, The Artificial Intelligence Community at MIT, Bourbaki, Lisp, Church's lambda calculus and Alan Kay's Dynabook Project (Papert, 1980, page 210). Logo is explicitly intended to appeal to the emotions as well as the intellect (Papert, 1980, page 7). One of Logo's chief proponents, Seymour Papert, sees computers as potential tools for exploring how thinking and learning work, and as possessing a double-edged power to change how people think and learn. (Papert, 1980, page 209) Logo has taken many forms, but the best known and best explored version is "turtle geometry" Logo.

Turtle geometry is a computational version of geometry which can be explored by giving commands to a "turtle" which leaves a trace as it moves. (The turtle may be a physical robot or may move about in two dimensions on computer terminal screen.) Commands can be assembled in packages called procedures, which can be re-used, adapted, or used as building blocks for further procedures. (Lest readers from non-Artificial Intelligence backgrounds think that Logo is only a toy for children, it should be stressed that while easy to use, Logo is both theoretically and practically a powerful programming language.) Strong claims are made for Logo by its proponents. Speaking in the context of the way in which the notions underlying differential calculus crop up naturally in Logo when trying to "teach" a turtle to move in a circle or spiral, Papert says that "I have seen elementary school children who understand clearly why differential equations are the natural form of laws of motion." (Papert, 1980, page 221). To attribute this understanding to Logo would be a very strong claim, (and Papert only makes it implicitly) since this is an understanding that escapes many graduate mathematicians, physicists and engineers. The validity of the claims made for Logo are hotly contested. Papert says that the sort of growth he is interested in fostering can take decades to bear fruit, and would

not necessarily show up on "pre and post" tests (Papert, 1980, page viii).

2.2 Bamberger's Music Logo: two key experiments and their implications

In one sense, any attempt to do, using music, what turtle geometry does using mathematics could be considered as a music Logo, But the term 'Music Logo' is usually identified with the inspiring research of Jeanne Bamberger and her colleagues, working in association with Seymour Papert, Terry Winograd, Hal Abelson and others at MIT at the time.

Music Logo has aims as ambitious as those of Turtle Logo. One aim is to describe music in a way that allows redefining of its elements by the user, and leads to extensible concepts. The representation should have very wide scope, yet be uniform and simple. It should allow any student to create her own tools to explore and create music that makes musical sense to him, whatever that sense may be. Music Logo should not only develop the skills that underlie intelligent listening, performance, analysis and composition, but should also encourage non-musical learning and understanding.

We now describe two experiments by Jeanne Bamberger that play a key role in her conception of a music Logo. (The following draws heavily on Bamberger 1972 and 1974.)

2.2.1 Tune Blocks

This experiment starts with a four voice synthesizer connected to a computer. A simple tune (say a nursery rhyme) has already been broken into motives or perhaps phrases, depending on the precise variant of the game being played. The phrases are labelled, say, B1 to B5 and if you type in B1, then the synthesizer plays the corresponding phrase. In one game, you are simply given a set of phrases, and invited to arrange them, using sequence and repetition, in an order that "makes sense" to you. The game sounds trivial, but strong claims are made for it. First of all, anyone can try it - there are virtually no prerequisites. Secondly, it involves active manipulation of and listening to musically meaningful elements (i.e. phrases

or motives as opposed to isolated notes or arbitrary fragments). Thirdly , it promotes "context-dependent" as opposed to isolated listening (e.g. block 1 followed by block 2 may give both blocks a new slant). Listening to what people say as they play the game, Bamberger reports that whereas a phrase is initially perceived in only one way e.g. "went down" or "went faster", as the game proceeds, it comes to be perceived in a variety of ways, for example "same downward movement as block 1, but doesn't sound like an ending". Players produce a wide variety of resulting pieces, and comparing other people's pieces, especially if they come from a different musical culture, is one of the pleasures of the game. Bamberger sees the use of neutral tags like B1, B2 as opposed to more visual representations as a positive advantage to tune blocks, as it means that visual sense cannot be used as a substitute for listening. There is a variation of the game in which the player is given a complete piece, programmed in advance, which she can hear at will, and the challenge this time is to recreate the original tune from its constituent phrases. Bamberger points out that in this version, the student is likely to discover - as an effortless by-product of play - the melody's overall structure (e.g. A B A). Players begin to ask themselves leading questions like "Why does B1, B3, B2" seem incomplete, although it sounds as though it has ended?". We will go on to look at the "Time Machine ", and Bamberger's suggestions for further work before trying to compare music Logo with its claims and aims.

2.2.2 The Time machine

Most people can clap along with both the rhythm of a melody and its underlying beat (some can tap both at once using hands and feet). The Time machine attempts to represent the interaction of these two streams of motion in a way at once intuitive and concretely representable. The gadgetry is very simple. The player strikes a drum, which leaves a trace on a VDU screen. Figs 6, 7 and 8 should make clear the relation between what is struck on the drum and what appears on the screen. Pieces can be recorded and played back later with synchronised sound and display. One crucial feature is that the time machine can play a regular beat against which you can input your piece; both the rhythm and the beat will appear on the screen (see fig 9) as you play. (Bamberger refers to the rhythm shown in fig 9 as "Mary had a little Lamb", but British readers might prefer to think of it as "London Bridge is falling

down".) Once recorded, the beat trace can be concealed while you listen to both streams of sound at once. Bamberger claims that one activity in particular - trying to picture how they mesh, and then revealing the beat trace to check your guess - can transform how people perceive pieces (Bamberger, 1974, page 13). Players often originally perceive the song shown in fig 9 both aurally and visually as consisting of three "chunks" - as in fig. 8.

Fig 6 Fast regular beat

Fig 7 Slow regular beat

Fig 8 "London Bridge is falling down"

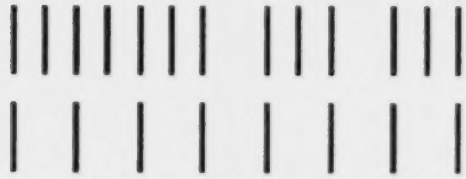


Fig 9 "London Bridge is falling down" (with pulse)

Fig 10 "London Bridge is falling down" (with pulse & linked trace)

Figures 6 - 10 redrawn from Bamberger (1974)

Bamberger shows that the single suggestion of *linking visually all the marks that occur during the lifetime of each beat* (fig. 10) can lead to the recreation of conventional rhythm notation, and can lead to the perception (though some people already perceive this anyway) that the first chunk contains a sub-chunk identical to

the second and third chunks. This is the only "game" that Bamberger describes for the Time Machine in the papers cited[1].

2.2.3 *Tackling higher level features of music*

Having considered two key games of Music Logo that deal with small scale musical structures, the question arises to what extent a similar approach could help explicate higher level structures in music. Given sufficiently extensible representations, it appears that musically sensible transformations could be carried out in a Music Logo at various levels of musical structure, and it could be argued that this is an important part of what composers do. Bamberger quotes from Schoenberg (1967):

> Even the writing of simple phrases involves the invention and use of motives, though perhaps unconsciously. . . The *motive* generally appears in a characteristic and impressive manner at the beginning of a piece. . . Inasmuch as almost every figure within a piece reveals some relationship to it, the basic motive is often considered the "germ" of the idea. . . However, everything depends upon its use. . . Everything depends upon its treatment and development. Accordingly variation requires changing some of the less important ones [features]. Preservation of rhythmic features effectively produces coherence... For the rest, determining which features are important depends on the compositional objective.
> (Schoenberg, 1967, page 8).

Bamberger mentions briefly experiments with slightly more extendable representations; in particular a language for representing and changing separately the pitch and durational aspects of a melody. But for more complex musical structures, Music Logo needs more extendable representations than those suggested so far, if it

[1] The use of the time machine links very interestingly with slightly later work by Bamberger (1975) in which she analyses different notations (some of them invented spontaneously) used by a cross-section of people - children, novices and trained musicians - to record rhythms. The notations are classified into two types, figural and metric, and it is claimed that they correspond to two ways of perceiving rhythm. It is also claimed that individuals tend to favour one perception or the other, but that both ways of perceiving must be learned as a pre-requisite for full musicianship.

is to permit easy musical manipulation. However, the implied claims made for the effects on users' perceptions of using even relatively crude tools like Tool Blocks and the Time Machine are interesting.

Bamberger reports how, for a set of students who had played the two games for several weeks, the experience of listening to Haydn's symphony no. 99 was transformed by a perception of the piece's evolution from a 5-note motive. The new perception, for some, dramatically extended their musical taste.

It is not hard to believe that better notations and visual representations can be found than conventional staff notation, and it is imaginable that theories of tonality, rhythm and harmony could be given computational embodiment and recast as discovery learning areas (more on this in part II of the thesis); but Music Logo has such ambitious aims that fulfilling them is likely to be very hard.

2.2 4 Music Logo the language

Terrapin Music Logo (Bamberger, 1986) is a version of Turtle Logo that "replaces graphics with music primitives". As in Turtle Logo, programs are sequences of function calls. The central data structures in Terrapin Music Logo are lists of integers representing sequences of pitches and durations. Pitch lists and duration lists can be manipulated separately before being sent by a command to a system buffer ready to be played by a synthesiser.

Note lists and duration lists may be manipulated using arbitrary arithmetic and list manipulation functions, and the language supports recursive programming. Random number generators are also available. List manipulation functions corresponding to common musical operations are provided. For example, one function takes a duration and pitch list and generates a specified number of repetitions of the phrase shifted at each repetition by some constant specified pitch increment - creating what musicians call a "sequence". Other "musical" functions and their effects include;

- retrograde - reverses a pitch lists,
- invert - processes a pitch list to the complementary values within an octave,
- frag - extracts sublists,
- fill - makes a list of all intermediate pitches between two specified pitches[2].

Bamberger suggests many simple exercises that are variations on the basic activity of iteratively manipulating the list representations to try to reproduce some previously heard musical result, or conversely making formal manipulations on lists and procedures and trying to guess the musical outcome. Bamberger stresses the importance of the "shocks" and learning experiences precipitated by two particular classes of phenomena; firstly, small manipulations of the duration list which produce radically changed perceptions of how the pitch list is chunked and secondly, the unpredictable disparity between degree of change in the data structure in general compared with the degrees of perceived change they produce.

2.3 Loco

One of the most elegant latterday Music Logos is Loco, developed by Peter Desain and Henkjan Honing (1986). Loco is a set of conventions and extensions to Logo for dealing with music composition and constitutes a microworld for specialised composition techniques - in particular the use of stochastic processes and context free music grammars. Desain and Honing make no claims that Loco embodies explicit knowledge about how to compose music - indeed, their whole philosophy is fiercely opposed to this approach. Desain and Honing argue instead that their flexible composition system should make the minimum restrictions on how the composer formulates her compositional ideas , and should make no assumptions at all about the instrument to be used and how it is controlled.

To this end, the bare minimum assumptions are built in to Loco about what instruments are to be controlled, and what commands each instrument will require. All that is assumed is that each command in the stream of commands to the

[2] Levitt (1981) makes extended use of the related, but more general concept of a 'melodic trajectory'.

instruments will take the form of a list or "object" containing a command name followed by zero or more arguments of some sort[3], and that some instrument will recognise and respond to each command. This gives the instrument-controlling end of the system great generality, and has allowed the same system to be used to control MIDI synthesisers, mechanical percussion installations, tape recorders, etcetera. The only other restriction on any instrument-controlling commands that the user may care to define is that a notional duration of the event that is triggered by any command must either be present as an argument to the command name or otherwise computable from the command. The net result is that Loco could in principle be used to control more or less any output device cleanly.

Loco has a very elegant, simple and general time structuring mechanism - without which all of the sound events would occur simultaneously. Two relations are provided, PARALLEL and SEQUENTIAL which can be used to combine arbitrary musical objects, and nested where appropriate. SEQUENTIAL is a function that causes musical objects in its argument list to be played one after another. PARALLEL is a function that causes its arguments to be played simultaneously. However, Desain and Honing are at pains to point out that before any data object is played, SEQUENTIAL and PARALLEL objects are treated as items of data that can themselves be computed and manipulated. The result of this framework is that arbitrary time structuring can be applied with a high degree of flexibility.

Moving on to tools for particular composition approaches, Desain and Honing provide two sets of primitives for composing using stochastic processes and context free grammars respectively. The key structuring technique used in both cases is that simple one line definitions are used to associate the name of a variable with a list of its possible values. In the stochastic microworld, depending on the choice of

[3] Desain and Honing have fitted their own object-oriented extension to LOGO, although not much detail is given about this aspect of the system. Because it is not reported in detail, it seems wise not to make too many assumptions about its nature. They refer to what we call instrument "commands" as musical "objects". However, to avoid making unwarranted assumptions, and to aid those readers unfamiliar with object-oriented approaches, it does not seem that any conceptual damage would be done by thinking of the word "object" in this context as meaning simply a list. In either case, perhaps the most important point is that any variable could be arranged to evaluate to an executable musical entity.

keyword used in the variable definition, subsequent evaluations of the variable may produce any of various possible effects, including;

- a random choice among its possible values,
- a choice weighted by a probability distribution,
- a random choice in which previously picked values cannot be picked again until all other possible values have been picked,
- selection of a value in a fixed order, returning to the beginning when all values have been picked .

The implementation technique in each case is that the "variable" is actually a simple program created by the "variable definition", but this is invisible to the naive user. Typing the name of the variable for evaluation simply produces an appropriate value each time in accordance with the one line definition.

Other one line "variable definitions" create variables that increment themselves on each evaluation by some amount specified in the definition. At this point, elegant use of the facility for functional composition provided by LOGO starts to give the system real power - the value of the increment could be specified as a stochastic variable, producing a variable that performs a brownian random walk. Brownian variables could then be used, for example as arguments in commands to instruments within some time structured framework. Techniques such as these are used to construct very concise, easy-to-read programs for transition nets, Markov chains and other stochastic processes whose precise operation can be modified using the full power of a general purpose programming language . In a similar way, rewrite rules for context-free (but no more powerful) grammars can be implemented almost directly, for example for rhythms. Competing rewrite rules can be assigned probabilities giving rise to a so-called "programmed grammar".

2.3.1 Criticisms of Loco

The primary design goals of LOCO include ease of use by non-programmers, modularity at all levels (to allow anything to be varied without damaging anything else) and user extendibility. The goals of modularity and user extendibility are well

met by the choice of a variant of LISP as the programming language, by the elegant conventions for time structuring and instrument control, and adherence to clean functional programming. Ease of use by non-programmers is helped by the use of the LOGO variant of LISP which avoids All Those Confusing Brackets (and seems to be attested by the successful use of LOCO in workshops for novices and professional composers). All of the primary goals are also helped by the elegant program generation techniques used in the composition microworlds. LOCO is useful for novice and experienced programmer alike not because it embodies any knowledge about music, but because it is a very well-designed programming language that is comparatively easy for novices to learn and which contains time structuring and instrument interfacing conventions that are modular, flexible and general.

What Loco could not do by itself, and what it does not aim or claim to do, is to teach in any explicit way how, for example, tonal harmony works, or what kinds of purposes rhythm serves, or what constraints are implicit in a certain style of music. Loco is useful for learning music, not because any musicality is embodied in it, but because it is an elegant, relatively easy to use programming tool of great modularity, flexibility and power.

3 Tools for the student

In this category of application of computers to the teaching of music, the computer is being used by the student as a tool to assist in carrying out some task. This contrasts loosely with both CAI in which the student is typically being managed, rather than being a manager, and also with music Logo, which has a stress on free exploration, rather than instrumental use. Three examples of tools are score editors, computer musical instruments and music analysis programs. As has already been mentioned, music editors and computer musical instruments are not included in this discussion; but it is useful to briefly summarise their most obvious advantages and disadvantages.

On the positive side: computer musical instruments (CMIs) allow students to compose and play pieces virtually unrestricted by their level of manual performing

skill; music editors allow students to listen to music scores, and hear the effect of any changes they may wish to make, in a way impossible with textbooks, records or tapes. On the negative side: firstly, ways of controlling computer musical instruments in musically meaningful ways are often highly restricted; and secondly, most music editors tend either to require knowledge of common music notation (which can be very difficult for beginners), or to only offer the most rudimentary structuring methods. Both of these problems will be returned to in the section on the musician-machine interface.

We now move on straight away to analytical tools. The following section draws mainly on work by Blombach (1981).

The most common uses of the computer in music analysis are essentially for event counting, sorting, pattern recognition and statistical analysis. Typical analysis programs can recognise occurrences of pitches, note values, intervals and combinations and sequences of these elements. One use of such a tool could be to test the validity of statements by music theorists about the piece being studied. For example, it is possible in studying the Bach chorales to:

> "determine the range of each of the four voices, or the number of times pairs of voices cross and compare the results with standard elementary theory textbook statements. Or they (the students) might use the computer to find occurrences of parallel perfect fifths in the Bach chorales or examine resolutions of tritones. . . . Students find these exercises especially satisfying if they prove the textbook author's discussions inaccurate, imprecise or incomplete" (Blombach, 1981, page 75).

The principal problem with this sort of analysis is that, as Dorothy Gross says (Gross, 1984), there are "serious questions regarding the significance of mere counts of events". Such counts may help in establishing, for example, authorship of pieces, but there is no prospect that event counting can distinguish good music from bad. Such analytical tools may help identify places in the great composers where, for example, part-writing rules have been observed or broken, but they do not get us

much closer to knowing the significance of these "rules", or when it makes musical sense to break them. Statistical analysis tools are limited in what they can teach because they connect neither with any developed theory nor established practice of music analysis. Later in the literature review, tools for musical analysis that reach a little deeper will be considered, but now we turn to the final section on current practice, about an educational music game with roots a little different from those of Music Logo.

4 Interactive graphics music games

Music Logo is by no means the only educational music game or the only music discovery learning environment. In this section, we describe an interactive graphical music game presented by Martin Lamb (1982). Although the game has many similarities with Music Logo (which Lamb mentions), its modest aims and roots in traditional music notation put it in a different (though not necessarily inferior) category. Lamb says that the purpose of the game is:

> 1/ to provide musically untrained children with a means of inventing their own compositions and hearing them immediately, and
>
> 2/ to teach musical transformations (e.g. augmentation, transposition and inversion) by letting children manipulate their visual analogues and immediately hear their effect.

The game is best described using pictures. The student begins with a set of blank staves (Fig 11). The player points to the word "Draw" on the display, using a pointing device, and then, holding down a button on the pointing device, can sketch a freehand curve on the staves (Fig 12). The program then straightens the curve to preserve left-to-right flow (Fig 13). Next, the program plays a line of music with the same shape (pitch for height) as the straightened curve. The curve then breaks up into discrete triangle shapes, each representing a note (Fig 14) (This diagram fails to show properly that a note can lie on a line or a space, or midway between the conventional positions, to indicate black notes. The length of each triangle is proportional to the length of the note.) Pointing to "transform", a group of notes can

be selected by circling them (Fig 15), and an inner circle can be drawn to exclude notes. A copy of the selected notes can then be moved about the screen (Fig 16). Using one of two physical sliders, the selected fragment can be continuously stretched or squashed relative to a horizontal axis (time relations are preserved).

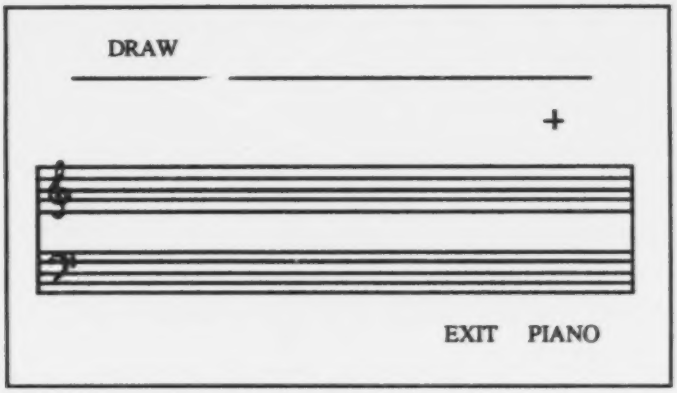


FIG 11 Lamb's Interactive graphics game: initial set up

FIG 12 player draws a curve

Figures redrawn from Lamb (1982)

The image can be squashed beyond a completely flat position continuously through to its mirror image. When the image reaches exact inversion, it indicates this by glowing more brightly. Similarly, using the other slider, a fragment can be squashed vertically into a chord, beyond simultaneity if desired to its exact retrograde and further (interval relations are preserved). The reaching of the exact retrograde point

is indicated when the mirror-image fragment glows more brightly. At any point, the transformation can be played at the pitch at which it is floating, or moved down to a different pitch level. It is possible to move or delete notes, as well as copy them, and there is an undo command.

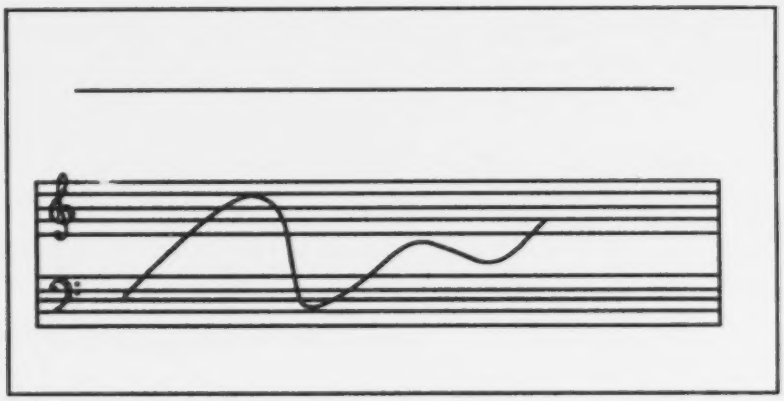


Fig 13 Curve is straightened to preserve left-to-right flow

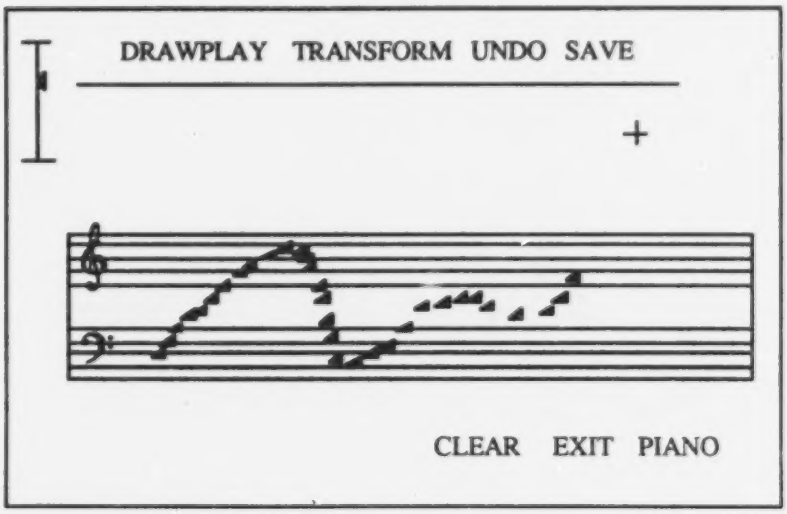


Fig 14 Computer changes curve into discrete notes

Figures redrawn from Lamb (1982)

It appears that more than one voice can be entered onto the screen by sketching more than one curve, and music can also be entered using a keyboard. The display uses a colour coding system. Different voices have different colours, but material derived

by transformation from the same material is displayed in a single colour. This game harmonises four elements ; pitch and time, their clear visual analogue, their gestural analogue and traditional staff notation. Not only are pitch and time represented in this way, but also the musical transformations transposition, augmentation and diminution, and the relations inversion and retrograde. This game appears to amply fulfill the modest objectives stated above. What is unclear is how far these objectives go towards developing composition skills.

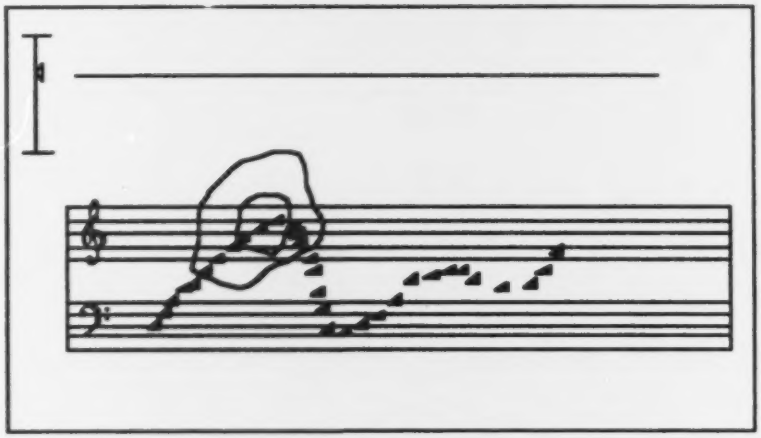


Fig 15 User has selected transform. Circled notes are to be transformed
Notes in inner circle are to be excluded

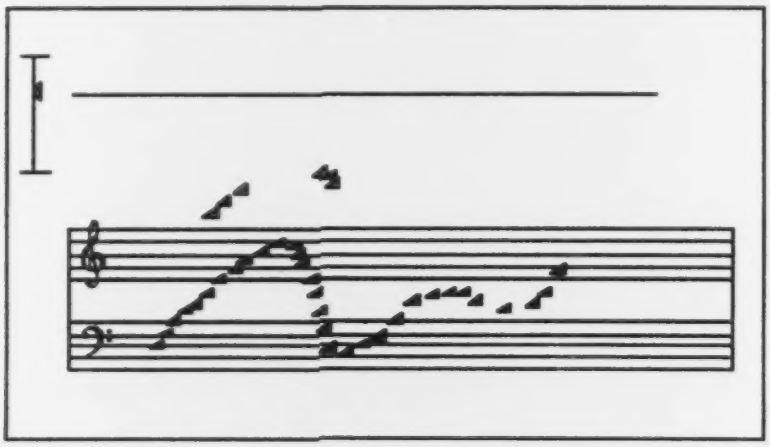


Fig 16 Notes to be transformed are displayed as hollow triangles

Figures redrawn from Lamb (1982)

5 Conclusions on early approaches

Computers have been used up to now in music education mostly as administrators of multiple choice questionnaires, as easily used musical instruments, "musical typewriters" and for musical games; they can count and identify simple musical events, and perform simple musical manipulations such as reordering, transposition, and pitch and time transformation.[4] High quality exploratory microworlds for music are starting to introduce a new element on the traditional scene.

None of the existing microworlds considered make available *explicit* computational representations of the kinds of musical knowledge, for example about harmonic interrelationships, that a novice may need when learning to compose tonal music.

A human teacher who had this knowledge already could use the microworlds to devise individually relevant experiments to help students gain the experience they need to be able to understand textbook descriptions of the knowledge. In the next section we discuss what more might be needed, where expert human help is scarce, to encourage and facilitate composition, and speculate how artificial intelligence might go some way to help provide it.

[4] Although this review has included material representative of most styles of approach, it has been highly selective. Some other work in the area that we have no space to pursue in detail is noted below. Distinguished pioneering work in the application of computers to music education has been carried out by Sterling Beckwith (1975). Well-known composers who have contributed to music education using computers include Xenakis (1985). An excellent set of simple ways for novices to explore musical ideas by themselves with a computer can be found in (Jones, 1984). Related work can be found in (Winsor, 1987). Music Logos have also been designed by Orlarey (1986) and Greenberg (1986).

Chapter 3: The musician-machine interface

The purpose of this chapter is to look at recent developments in the 'musician-machine interface' that are likely to affect the future use of computers in teaching music. We begin by reviewing some general reasons for the importance of the musician-machine interface in music education and consider some general strategies for designing good interfaces. In the subsequent sections, the following topics are covered: we consider connections between research in the musician-machine interface (MMI) and research in human-computer interaction (HCI) in general; we examine some possible abstract relationships between musician, score and instrument; finally we make a brief selective survey of three contrasting types of musician-machine interface with educational relevance and draw conclusions. The structure of the chapter can be summarised as follows;

1 The importance of the interface
2 General strategies for designing good interfaces
3 MMI meets HCI - a two way traffic
4 Intelligent Instruments: new relationships between musician and machine
5 MMI in music education
- 5.1 Use of common music notation in interfaces.
- 5.2 Interfaces using graphical representations of structure
- 5.3 Control panels for computational models
  - 5.3.1 Desain's proposed graphical programming tool.
  - 5.3.2 Hookup
  - 5.3.3 Comments on Hookup and Desain's proposal
  - 5.3.4 The 'M' program
  - 5.3.5 Jam Factory

6 Conclusions on new developments in the musician-machine interface

1 The importance of the interface

Early computer interfaces mimicked typewriters, and so had to organise their input

and output as strings of words. For many tasks, this is a highly unsuitable form of communication. Buxton (1987) points out that few motorcyclists would be happy driving around central London in the rush hour using an interface based on strings of words to control their bike. In many respects, education in music composition is a task equally ill-suited to be carried out through the medium of text-based interfaces. To mention but one disadvantage, recent research has suggested that the very gestures involved in performing music can play vital roles in the process of music composition (Baily, 1985). But can we conclude that interface design problems in specialised domains can be solved merely by designing and employing appropriate input/output hardware for the task? Such a conclusion would be seductive, bearing in mind that a host of new and subtle input peripherals such as midi-gloves[1] in principle allow computers to be controlled in arbitrary respects by arbitrary gestures. Unfortunately, human-computer interface design problems run much deeper. The use of arbitrary gestures to control arbitrary parameters tends, in general, to produce arbitrary interfaces. Another underlying problem is that it may be hard for a novice to exercise appropriate control of a complex process without extensive knowledge or experience of the process being controlled.

2 General strategies for designing good interfaces for novices

One general strategy for designing interfaces for novices, aimed to get around the last mentioned problem, is to represent explicitly and perspicaciously in the interface the implicit knowledge required to operate in the domain. Another general strategy for designing interfaces for users who lack relevant knowledge or experience is to exploit the novices' existing knowledge or propensities (Smith, Irby, Kimball, Verplank and Harslem, 1982). When using either strategy, it is often helpful to study the cognitive psychology of the relevant domain in advance: ideally, both the structure of the domain *and* relevant human propensities should be taken into account when devising mappings of gesture, feedback and action to each other for

[1] As the name suggests, midi-gloves are hand-worn computer peripherals (demonstrated at the 1986 International Computer Music Conference), that allow finger and hand gestures to generate Midi (musical instrument digital interface) signals and thus in principle control arbitrary computer actions in real time. Midi gloves are related to work by Tom Zimmerman, Jaron Lanier and others (Zimmerman, Lanier, Blanchard, Bryson, and Harvill, 1987) in developing a general purpose hand gesture interface device, the 'Data Glove' (which also gives tactile feedback).

the interface.

3 MMI meets HCI - a two way traffic

Before looking at issues that arise most acutely in music, we will briefly consider developments in human-machine interfaces in general. As we noted in passing at the beginning of the section, until very recently, computer interfaces in all fields were largely text-based, typewriter like and designed around the convenience of the designers and manufacturers. These interfaces required long-term memorisation of extensive command languages and syntaxes and short term memorisation of the current state of the system. Work at Xerox PARC (Smith, Irby, Kimball, Verplank and Harslem, 1982) and elsewhere led to a wave of radically different interfaces. The new interfaces were designed with the help of psychologists for the convenience of the user and made heavy use of graphics and simple pointing. The new interfaces lowered cognitive load by exploiting novices' previous knowledge through visual metaphors and displaying the state of the system graphically. The shift in interface design methodology involved many other detailed changes in design philosophy that we do not have space to explore. These changes affected interface design in computer music, but it is also true to say that, due to the demanding nature of interface design problems in computer music, there has been considerable traffic of innovations in the other direction. Buxton (1987) argues that research in computer music interfaces in many respects anticipated (Buxton et al., 1979), and in some cases influenced mainstream interface research in areas such as menu design. Buxton argues further that present day developments in the musician-machine interface, such as use of multiple pointing devices, are once again anticipating developments in mainstream human-computer interaction. Whatever the truth of this, a central point about human-computer interface design is that the deep issues no longer simply concern matters of physical ergonomics; good interface design for complex cognitive tasks is nowadays often bound up with the construction of good cognitive models of the task. Hence because of the complexity of the cognitive skills associated with music; its multi-modality, multi-media demands; and because of its intertwined, nearly decomposable, temporal dimensions; music can be seen as a demanding *forcing function* in computer interface design in general.

4 Intelligent Instruments: possible relationships between musician and machine

Before considering some particular musician-machine interfaces, it may be helpful to consider the musician-machine interface at a more abstract level. Mathews and Abbott (1980) define three different abstract relationships between musician, score and instrument - two familiar and one relatively new. The relationships can be described pictorially (Figures 17-19). Figure 17 illustrates the way that;

> ". . with traditional instruments, all the information in the score passes through the musician, who communicates the information to the instrument via physical gestures. The musician has fast, absolute control over the sound, but he or she must be able to read the score rapidly and make complex gestures quickly and accurately." (Mathews with (sic) Abbott, 1980, page 45)

The second relationship is that which holds with musique concrete, player pianos and music programming languages (Fig 18). The score completely specifies the performance, and although this can be amended indefinitely off-line, no real-time demands or opportunities fall to the player during performance. Fig 19 shows a third possible relationship. The entire score does not have to be interpreted in real time by the musician, but the musician can in principle reserve any aspects of the music for his interpretation.

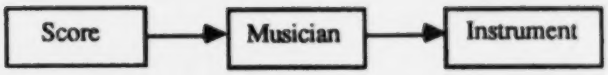


Fig. 17

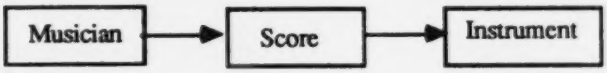


Fig. 18

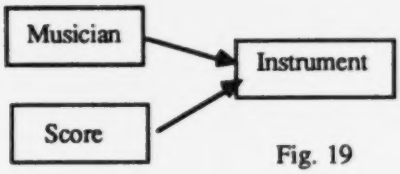


Fig. 19

Musician/score/instrument relationships 1, 2 and 3 respectively
Figures redrawn from Mathews with (sic) Abbott (1980)

> "The score is a sequence of *partial descriptions* of musical events. In general the sequence is ordered by time. The information not contained in the score is supplied by the musician in real-time during the performance. The aspects of the performance that are to be interpreted are generally supplied by the musician.The aspects not subject to interpretation are supplied by the score." (Mathews with Abbott, 1980, page 45)

A special case of this third relationship is of course, the familiar conductor-orchestra relationship. No representation and control methods are known at present, needless to say, that could allow a human-machine relationship to approach the subtlety of the orchestra-conductor relationship. However, some special cases of relationship 3 may be very useful despite their relative lack of sophistication. One of the most obvious is to separate rhythmic and pitch information, as was done by Bamberger. Mathew's "sequential drum" can go some way beyond this. The drum is a rectangular surface that can be played by hand or stick, and has four outputs; hit time, hit strength, and X and Y position on the drum's surface.The four signals can be used to control any synthesizer parameters; loudness, timbre, location of sound, time of occurrence, or perhaps even the tempo of groups of notes. The score might,

in a simple case, contain a rhythmless melody, allowing the musician to concentrate on phrasing and accent. At least one composer, Joel Chadabe (Chadabe, 1983) performs, or "interactively composes" - his term - in much this way, although Chadabe uses two theremin rods [2] rather than sequential drums as controllers. The boundary between musical performance and composition has always been blurred, but the developments we have been discussing may make the two activities even harder to distinguish in the future.

It might be wondered if this discussion of musician-machine interfaces has much to do with artificial intelligence and learning composition skills. There is a strong argument that it has. Artificial intelligence research has long been inextricably intertwined with efforts to improve the human machine interface in a wider context. As we have already noted, musicians' physical gestures in making music may be an essential element in a description of what music is about (Baily, 1985). Conversely, musical intelligence may become a useful component in responsive musical instruments.

5 MMI in music education

We now turn to developments in the musician-machine interface that have special potential relevance to educational concerns. A good survey of the musician-machine interface in general can be found in Pennycook (1985a) and a survey of the role of graphics in computer-aided music instruction can be found in Pennycook (1985b and 1985c). Human computer interfaces for music are used for a wide variety of tasks, and it can be hard to draw clear distinctions between those with and without educational implications. Typical tasks addressed by existing interfaces include score editing, real-time performance, assisting composition, computer-aided instruction, signal processing, assisting analysis, computer musical instruments, manipulating recorded sound and others, all of which can overlap to a greater or lesser extent. A better way of classifying the various approaches than by the ostensible purpose for which the interfaces are designed is to consider the main representation or representations offered to the user by the interface for visualisation or control.

[2] Theremin rods are proximity-sensitive antennae used to control electronic musical instruments by fluid hand gestures, invented by Lev Theremin in 1920.

Recurring representations include textual strings; midi streams; sound input and output; representations of key or finger boards; common music notation; two and three dimensional graphs and arrays; trees and nets; and control panels for computational models. We will discuss only three classes of representation in interfaces, namely common music notation, graphical representations and control panels for computational models. One reason for picking these particular classes of interface representation for discussion is that they are particularly suitable for communicating high-level structure. We will begin with the use of common music notation in computer interfaces.

5.1 Use of common music notation in computer interfaces.

Common music notation (CMN), with origins traceable back at least as far as the 11th century and music educator Guido D'Arrezzo, is a highly sophisticated notation that makes use of many clever devices to make musical structure visible and reduce cognitive load by chunking. Ingenious notational devices include the key signature, visual grouping by meter, cartesian encoding of time and pitch, and visual gestalts for certain aspects of scales and chords. Direct manipulation of notes on a graphic score is now a ubiquitous interface technique in computer music. Many of the most sophisticated ideas for the use of CMN in computer interfaces now in widespread use were first developed by Buxton et al. (1979) as part of the Structured Sound Synthesis Project at the University of Toronto. This project was notable as the first major *systematic* investigation of human machine interface design issues in computer music. Pennycook (1985a), notes that this project drew on earlier work on the musician-machine interface including: Pulfer (1971) and Tanner (1972) and the "Computer Aid for Musical Composers" project at the National Research Council of Canada; Vercoe (1975) at the MIT Experimental Music Studio; Truax's POD system (1977), and work by Otto Laske (1978).

The sophistication and ubiquity of common music notation make it an excellent starting point for designing interfaces for those with considerable musical experience and training. However, common music notation can act as a formidable barrier to the novice. It can take a substantial effort for a novice to learn to read music fluently

and with understanding: typically a few years practice and study is required. One reason for the difficulty is as follows: music notation chunks some concepts (such as the concept of key) very compactly, but as a result it may take prolonged effort for the novice to come to understand what is only *implicit* in the notation. A few examples are given below of the way that common music notation calls on substantial implicit knowledge for its decoding (many more examples could be given).

- The stave notation makes pitch distances explicit in terms of degrees of the scale, but obscures them in terms of chromatic steps: the distance in semitones between two points on the stave depends on the key signature in force.

- The structure of the diatonic scale as a whole (in terms of semitone step sizes) is not shown anywhere explicitly in the notation.

- The varying quality of scale-tone chords is not made explicit by the notation. The inference of chord quality from the notation depends on experience (or calculation based on abstract knowledge) of key structure. So, for example, major, minor and diminished triads may appear visually indistinguishable from each other in the notation.

- Tonal centres are not indicated explicitly, their location must be calculated using memorised abstract rules.

- The rules for constructing key signatures require use of a memorised abstract procedure.

- Relationships between notes in remote keys are difficult to decode visually.

In summary, the same ingenious chunking of musical concepts that helps to make common music notation so useful for experienced musicians can make it an obstacle for novices.[3] For this reason, while there will always be an important place for

[3] Lamb's (1982) interface, reviewed earlier, is an interesting apparent exception to this claim, but

common music notation in the musician-machine interface for seasoned musicians, there is a need for other kinds of interfaces to help beginners. Such interfaces could help novices, especially those without good teachers, to gather sufficient musical experience, knowledge, vocabulary and motivation to begin performing a range of musical tasks with confidence, and to learn enough to be able to understand the basis of more complex notations. We will return to this theme in part II of the thesis.

5.2 Interfaces using graphical representations of structure

The second of the three classes of musician-machine interface that we will consider involves use of specialised graphic representations. We will deal briefly and selectively with just one representation of this kind, namely *tree-like representations*. (Other graphic representations such as *arrays*, and interfaces based on them, such as Music Mouse (Spiegel, 1986), will be considered in part II of the thesis.) Tree-like notations for music can appear in music interfaces in a number of guises; as a structuring device for scores (Buxton et al., 1979), as part of a notation for analysis (Smoliar, 1980) as a hierarchical structuring device for composition (Lerdahl and Potard, 1985) and as a notation for grammars. Tree-like notations are often used in collaboration with common music notation (Buxton et al., 1979). One interesting implicit use of tree-like structures in an unusual modality takes the form of aurally presented analyses of pieces of music in terms of shorter pieces of music. The key idea behind this kind of analysis is that some notes of the original piece are deemed less important than others and left out, revealing a series of recursively nested skeletal structures. (See Lerdahl and Jackendoff (1983) for such a theory.) Though interfaces based on trees are of great interest, and though there is no reason why such interfaces should not be designed for novices, at present most interfaces with a heavy commitment to this representation appear to be suitable only for users who already have very sophisticated knowledge or experience of tonal music.

5.3 Interactive control panels for computational models

The final of the three classes of interfaces that we will comment on is the class

on closer examination the representation used by Lamb bears only a very superficial resemblance to common music notation.

consisting of control panels for computational models. In one sense, of course, any programming language that can control musical peripherals, such as Music Logo or Loco constitutes such an "interface". Our focus, however, will be more particularly on *graphical* front ends to computational models. Note that a graphical front end to a programming language does not in itself constitute a computational model: a language is a *kit* for constructing such models. Some programs, on the other hand, offer graphical interfaces to some pre-designed, fixed computational model of some musical process. Yet other programs allow graphical tools to be used to interconnect large prefabricated components to assemble customised versions of some basic computational model. We will consider aspects of four interfaces that offer interactive control panels for building or controlling computational models of musical processes: two offer graphical interfaces to general programming kits, and two offering elegant graphical interfaces to processes that are in some way 'prefabricated' to a higher degree. The interfaces are: Desain's (1986) proposed graphical programming tool, Hookup (Sloane, Levitt, Travers, So and Cavero, 1986), 'M' (Zicarelli, 1987) and Jam factory (Zicharelli, 1987), which we shall now deal with in that order[4].

5.3.1 Desain's proposed graphical programming tool.

Desain (1986) describes an interface that allows a user to manipulate icons representing signal processing components (such as microphones, loudspeakers, amplifiers and delays) to construct signal processing devices on screen. The two essential software components of such a system are a graph editor and an incremental compiler for signal processing. Such components exist separately, but are not connected in Desain's system. Desain notes that if the graph editor were to be connected to a signal processing compiler and appropriate hardware, the user could test out the 'devices' she had assembled using a real microphone and loudspeaker. So far, this proposal relates directly to signal processing rather than music composition, but Desain goes on to propose that similar techniques could be used for programming a graphical front end to the Loco language, described earlier. This proposal seems entirely feasible, using either a graphic metaphor for functional

[4] See also Miller Puckette's excellent 'Patcher' (Puckette, 1988). Patcher is an iconic object-oriented programming language with MIDI (and other real-time control) facilities. Patcher is part of the Max environment associated with IRCAM.

composition, (as used, for example, in 'Icon machine Logo', discussed in Feurzeig (1986) for the domain of algebra), or using a graphical metaphor for message passing, as proposed by Desain. We will defer commenting on this proposal until after we have considered Sloane, Levitt, Travers, So and Cavero's (1986) related interface "Hookup". After this, we will comment on both interfaces together.

5.3.2 Hookup.

Hookup (Sloane, Levitt, Travers, So and Cavero, 1986) is a prototype software toolkit that allows icons to be interconnected on screen to create little working data-flow programs. Figure 20 shows a blank working screen. In the data-flow model of computation, quantities (typically numerical) 'flow' in one direction through a series of interconnected primitive processors, each processor producing a new output asynchronously whenever it is presented with a full set of new inputs. While waiting for a quantity to arrive at one input of a processor, quantities may be 'queued' on another input. The data flow paradigm of computing is normally used for signal processing, simulation and computer music applications (Desain,1986). In the case of Hookup, icons are provided for input devices that include the mouse itself, graphic sliders, buttons, clocks and boxes for inputting numerical quantities. Processors include arithmetic and logic components. Output devices include graphic 'sprites', graphs and midi output. Icons for some of these devices waiting to be connected up are shown in Figure 21. The specialised input and output devices provided by Hookup mean that it is suitable for animation and simple music composition. For example, Figure 22 shows a program that allows sliders and buttons to be manipulated to produce chord progressions. Other dataflow programs assembled in Hookup have realised, for example 'process music' compositions by the minimalist composer Steve Reich. We will now comment on Desain's proposal and Hookup together.

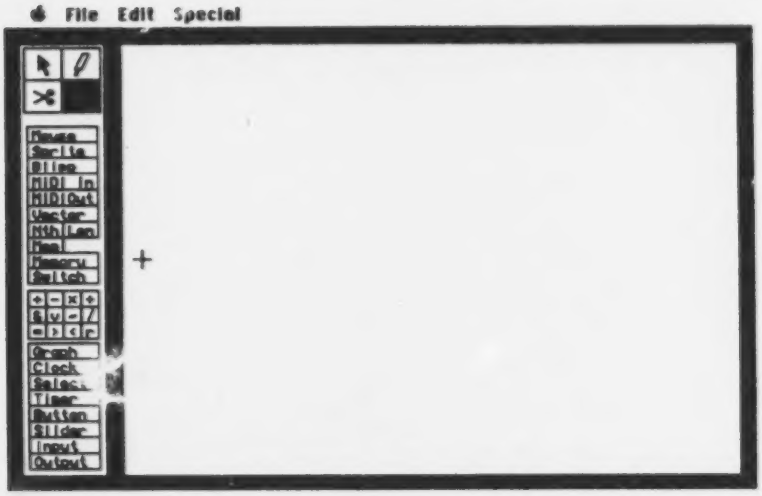


Fig 20 Blank Working Screen for Hookup

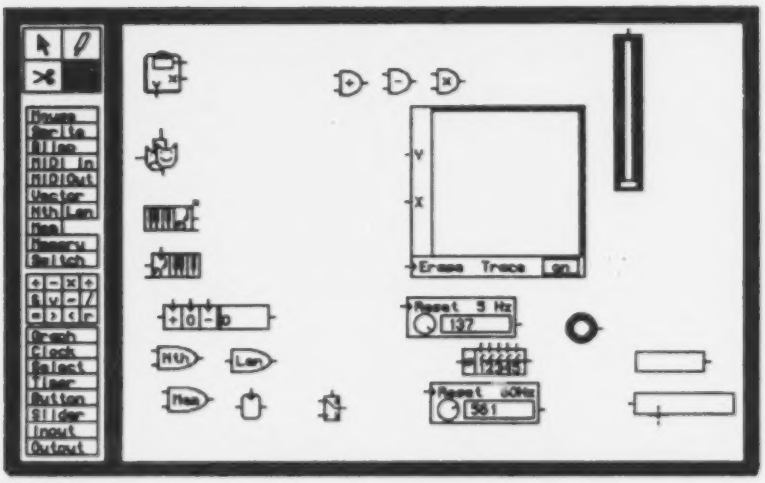


Fig 21 Hookup with various icons waiting to be hooked up

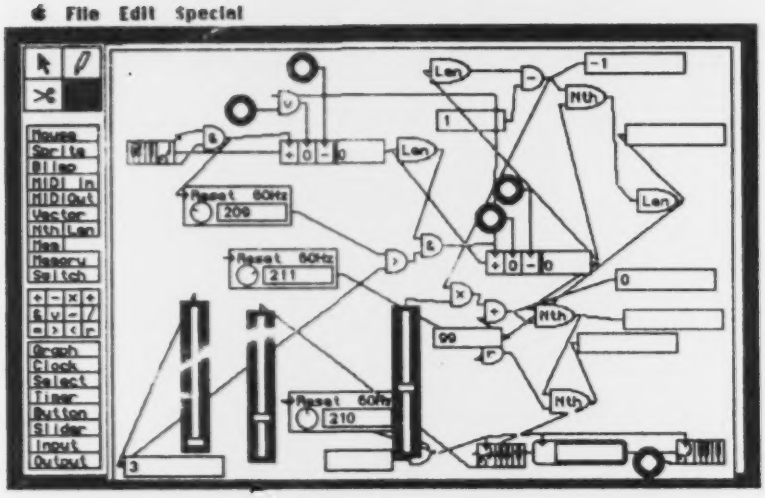


Fig 22 A Hookup program to play chord progressions

5.3.3 Conclusions on Hookup and Desain's proposed graphical front end to Loco

Both Hookup and Desain's proposed graphical front end to Loco are joyous ideas, full of promise, that could hardly fail to capture the enthusiasm and participation of novice would-be musicians. Both are outstanding contributions. However, neither remarkable program fully addresses the particular issues that we are pursuing, for the following reasons. Hookup is perhaps most accurately thought of as a tool for animation and Midi processing rather than than as a tool for novices to explore music composition (perhaps with the exception of 'process music'). Midi command streams manipulated by dataflow appear to be too low a level of description for modelling most aspects of tonal music perspicaciously. However, experiments on the use of Hookup to teach novices about aspects of music composition would be of the greatest interest. Desain's proposal, based on the full flexibility and symbolic power of Lisp, almost certainly has the potential to constitute an interface well suited for teaching novices about music composition. But while such an interface might constitute a good tool box for building computational models of aspects of musical expertise for novices to play with, it would not in itself constitute such a model. Given that one of our main pursuits is the provision of novices with explicit models of musical relationships and musical expertise in manipulable, perspicacious form, graphical front ends to general programming languages do not in themselves

provide us with what we require, and we need to look further. Now that we have looked at two general programming kits with graphical interfaces, we shall go on to look at our final two example interfaces, which offer elegant graphical interfaces to computational processes that are in some way 'prefabricated' to a higher degree.

5.3.4 The 'M' program.

M and Jam Factory are proprietary programs for the Apple Macintosh. (Zicharelli, 1987). M (Figure 23) allows sequences of notes, chords, rhythmic patterns, accent patterns and so forth to be built up in libraries and edited graphically. The sequences can be initially created using midi input from an instrument, or by means of various graphic editors. Particular note, rhythm and other patterns can then be graphically selected for combination to produce musical output, in real time, in up to four voices. Patterns can also be selected for real-time scrambling at random, or scrambling according to probability distributions in various ways. The result of changing which patterns are contributing to the output can be heard instantly. The scrambling can provide a degree of variation even when the user is not intervening.

Rather like the intelligent drum, M has a control area whose x and y value can be used to control which of the patterns in each musical dimension (pitch, rhythm, accent, etc.) are selected at any time to contribute to the live result. The choice of pattern in any of the musical dimensions addressed can be set to be controlled either by current x or current y value , or disconnected from the xy controller altogether. Either x or y position can also be set to control parameters other than temporal patterns, such as timbre and tempo. The links between parameters and controller can easily be changed in real time. As it happens, all of the musical parameters controlled have just six discrete values to chose from at any time (i.e. 6 note patterns, 6 rhythm patterns, 6 timbres, etcetera) except for tempo, which can be varied continuously. The selection of the current x and y values is done by moving an iconic baton around in the control area. By moving around the baton, the user can navigate between combinations of parameters and so "compose" interactively. M uses a variety of ingenious graphic input tools to facilitate the use of the various different editors, and uses carefully designed graphics to: represent the various patterns; show

what values are selected; and show what other choices are available for any parameter.

In all, M is well designed for "interactive composition". Most of the skill in using M to good effect seems to lie in the initial (and time-consuming) selection and creation of raw material to be mixed. M is a good tool to allow experienced composers, who know what they are doing, to experiment fluidly and flexibly with different combinations of their basic materials. However, M does not appear to address the particular issues that we are pursuing, in the following senses. Although novices can undoubtedly produce interesting and varied music with M simply by moving a mouse around about, there is little evidence that their newly found "skills" would make them any more able to compose when deprived of M, or able to articulate in any detail what they had been doing. With very careful selection of initial materials and structured guidance, novices might be led to experience educational "shocks" of the kind described by Bamberger (as reported the previous chapter). M is not in the business of providing novices with explicit models of musical relationships and musical expertise in manipulable, perspicacious form, except in the important regard of allowing pitch lists, rhythm lists, etcetera, to be combined perspicaciously. We will now consider the final interface, "Jam Factory".

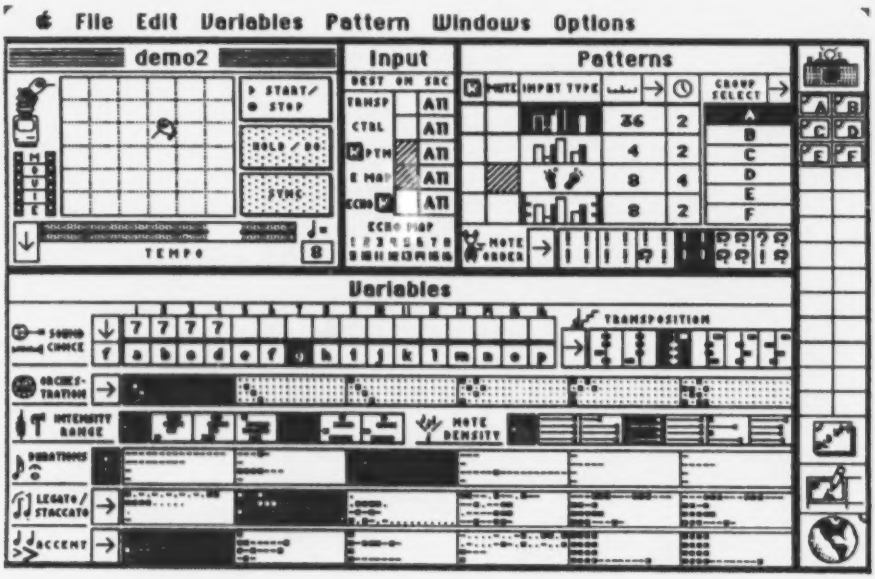


Fig. 23 A typical 'M' screen

5.3.5 Jam Factory

Jam factory (Figure 23) contains four "players" that "listen" to Midi input and "perform" instant variations on the input in real time. Thus, Jam Factory provides a sense of improvising or 'jamming' with the user. The variations on the original input are introduced by making probabilistic changes to the pitches and note durations of the input based on what has happened before according to information held in transition tables (Xenakis, 1971) associated with each player. The parameters of the transition tables (held separately for pitch and duration) can be varied by the user. The user may also introduce combinations of various kinds of deterministic processing into the changes made by each player. The deterministic variations affect such parameters as microtiming, articulation and the sustaining of notes that are followed by randomly introduced silences. Jam Factory has a well-designed (if initially confusing) interface and as a proprietary product is unique of its kind. From the point of view of our restricted viewpoint, the educational possibilities of Jam factory as regards composition are hard to evaluate. Jam Factory might act as a sort of 'Rorshach test' for experienced composers, echoing back to them variations of what they have played, from which they might pick out interesting variations for further development. In the case of novice would-be composers, leaving aside any general motivational benefits, it is hard to see any mechanisms by which novices exposed to Jam Factory might come any nearer to knowing how to compose using any other instrument or tool. Similarly, it seems unlikely that Jam factory would do much to help novices find a vocabulary to articulate musical experience. Needless to say, no criticism is intended of Jam Factory in its own right; we are merely commenting on its suitability for our particular aims. In general, one of the problems about probabilistic models of musical processes is that they tend to be superficial and crude.

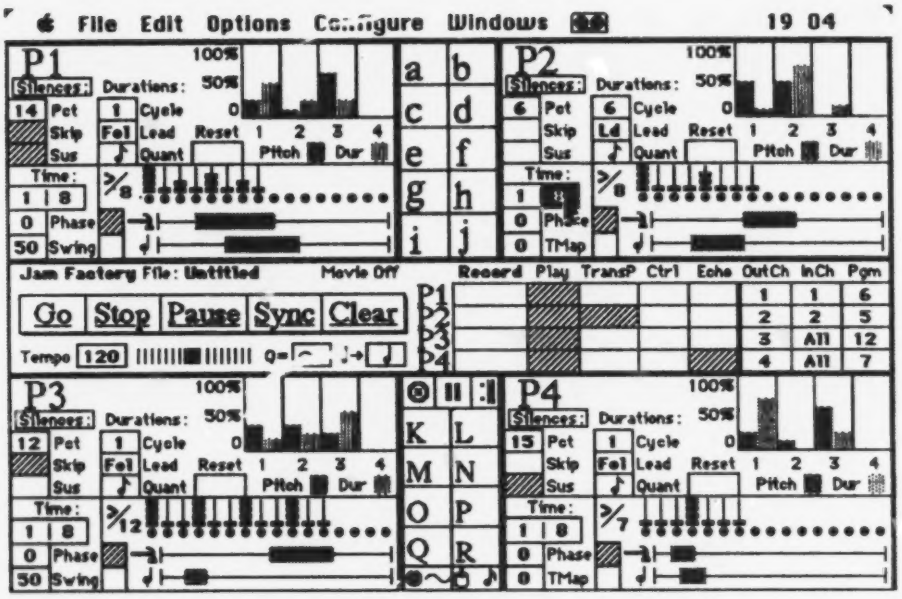


Fig. 24 Jam Factory

This concludes our comments on interfaces based on control panels for computational models, and our comments on contrasting types of musician-machine interface with educational relevance.

6 Conclusions on the musician-machine interface

We will now summarise the conclusions drawn from the survey of the musician-machine interface (in some cases repeating forms of words used earlier). It is hard to exercise appropriate control of a process without knowledge or experience of what is being controlled. One general strategy for designing interfaces to get around novices' ignorance or lack of experience is to represent explicitly and perspicaciously in the interface the implicit knowledge required to operate in the domain. Another useful general design strategy is to exploit users' existing knowledge or propensities (Smith, Irby, Kimball, Verplank and Harslem, 1982). When using either strategy, it is often helpful to study the cognitive psychology of the relevant domain in advance: ideally, both the structure of the domain *and* relevant human propensities should be taken into account when devising mappings of gesture, feedback and action to each other for the interface.

In the case of experienced musicians working in tonal music, the use of common music notation as the basis of interfaces has many benefits, because it allows sophisticated users to make full use of established chunking and structuring tendencies. In the case of novice musicians who want to learn to compose, its use as a sole or chief notation may often be inappropriate, because it may take novices several years to learn to read the notation fluently, and much important information is only represented implicitly in the notation. There is a need for other kinds of interfaces to help beginners. Such interfaces could help novices, especially those without good teachers, to gather sufficient musical experience, knowledge, vocabulary and motivation to begin performing a range of musical tasks with confidence, and to learn enough to be able to understand the basis of more complex notations.

In the case of experienced composers, a good design consideration (Desain and Honing, 1986) is to design interfaces that have maximum flexibility and make as few assumptions as possible about how a composer will work. This approach is consonant with the observed diversity of the way in which composers work (Sloboda, 1985). However, if we wish to educate novices in some particular aspect of music, it may be useful to reverse this design consideration and build into interfaces a very specific metaphor for some aspect of music.

Desain's proposal, Hookup, M and Jam Factory are all outstanding innovative graphic interfaces for manipulating musical materials. Diverse opportunities remain for constructing interfaces to provide novices with explicit models of various kinds of musical relationship and musical expertise at different levels in manipulable, explicit, form.

Chapter 4: Intelligent tutoring systems for music composition

The purpose of this chapter is to discuss the design, function and limitations of two existing intelligent tutors for music composition. We begin the chapter by noting the 'traditional' model for an intelligent tutoring system (ITS), and observe restrictions on the kind of domain in which this model is directly applicable. We then identify a restricted area of music composition likely to be amenable to this kind of treatment, namely harmonisation. Next, we review two intelligent tutors that teach particular aspects of music related to harmonisation. To help clarify the nature of the tutors, we briefly outline the goals, problems and methods associated with writing harmony. Finally, we draw conclusions about limits to the applicability of the traditional ITS model to music composition. The overall purpose of the chapter is to prepare the ground for work to be presented in part III of the thesis on a new framework for knowledge-based tutoring in creative domains. The structure of the chapter is as follows;

1 A role for the 'traditional' ITS model in music composition

2 Vivace: an expert system for harmonisation

- 2.1 The goal of four-part harmonisation
- 2.2 Problems inherent in harmonisation
- 2.3 How to harmonise a chorale melody
- 2.4 Pedagogic use of Vivace
- 2.5 Contribution of Vivace

3 MacVoice: a critic for voice-leading

- 3.1 Function and use of MacVoice
- 3.2 Criticisms of MacVoice and possibilities for further work

4 Lasso: an intelligent tutoring system for 16th century counterpoint

5 Conclusions on intelligent tutors for music composition

6 Related work in other fields

7 Conclusions on the use of computers in teaching music

1 A role for the 'traditional' ITS model in music composition

The first requirement for an intelligent tutoring system for any task is that it has to have some ability to perform, or at least to discuss articulately, the task in hand. This demands explicit knowledge of the task. Two other features are usually associated with ITS's: intelligent tutoring systems are expected to build up knowledge of what a particular student knows, so that opportunities can be seized to get points across in the most appropriate way or to diagnose misconceptions; secondly an intelligent tutoring system is expected to have explicit knowledge of ways of teaching (Self, 1974). The traditional model of an intelligent tutoring system may be entering a period of transition (Self, 1988), but it provides a good starting point. The very first requirement of the model, for explicit knowledge of the task, raises the question "for what areas of music composition do we have explicit knowledge ?". This appears to narrow the possibilities for this style of ITS contribution to music composition down to a very few areas. It seems entirely reasonable to suggest that intelligent tutoring systems might be built for factual aspects of music theory related to composition. These might not differ all that much from intelligent tutoring systems for, say, Geography (Carbonell, 1970). As regards the process of composition itself, one of the very few areas for which detailed, explicit rules of thumb can be found in textbooks is harmonisation. At least two intelligent tutoring systems for tasks related to harmonisation have been constructed, and we will now discuss their design, function and limitations.

2 Vivace: an expert system for harmonisation

Vivace is a rule-based expert system for the task of four-part chorale writing, created by Thomas (1985). Vivace has been used for teaching in ways that we will describe shortly. Vivace takes as input an eighteenth century chorale melody (i.e, roughly speaking, a hymn-like melody) and writes a bass line and two inner voices (tenor and alto) that fit the melody. We cannot sensibly discuss Vivace without some knowledge of what is involved in harmonisation. Accordingly, we will outline in the next three paragraphs what the task entails. (Experienced musicians might wish to skim from here to sub-section 2.4.)

There is no single correct harmonisation of a given melody, in fact the space of acceptable harmonisations of a given melody is very large indeed. There are a number of rules and guidelines given in text books for carrying out the task, abstracted from the practice of past composers. The rules and guidelines apply not only to each voice in isolation: but also to the sequence of chords that they all interact to produce; to interactions between any two pairs of voices; and to vertical spacing of the notes at any given time. There are also guidelines for cadences (chords that act as 'punctuation' to the harmonic sequence), guidelines for harmonic tempi, non-harmonic tones and various other considerations. Broadly speaking, the rules and guidelines are of four types. There are rules that act as firm requirements, rules that are statements of preference but not requirements, rules that are firm prohibitions and rules that prefer some occurrence to be avoided.

2.1 The goal behind four-part harmonisation

In essence, the aim of the rules appears to be to minimise the number of features that do not sound pleasing, and to maximise the number of features that do sound pleasing in all of the relevant dimensions (e.g. harmonic progression, voice-leading, interest of inner voices etc.). There are other genres (e.g. bebop, punk, minimalism) for which other overall goals seem to apply (e.g., induce surprise or shock; experiment with perceptually slow rates of change, etcetera). The simple goal we have outlined seems to characterise four-part harmonisation at a high level surprisingly well. One of the main problems in pursuing the aim is that setting up a pleasing feature in one dimension often destructively interferes with a feature in another dimension.

2.2 Problems inherent in the harmonisation task

We will now outline three of the chief problems inherent in the task of harmonisation. The first problem is that it is quite possible, indeed very common in beginners' classes, to satisfy all of the formal rules and produce a piece of music which is correct but is aesthetically quite unsatisfactory. The second problem is that most of the guidelines are prohibitions rather than positive suggestions. Milton

Babbit observes that "..tne rules of counterpoint are not intended to tell you *what* to do, but what *not* to do" (Pierce, 1983, page 5). The same applies to the closely related task of four-part harmonisation. To put it another way, viewing harmonisation as a typical AI 'generate and test' task, the rules constitute help in the testing rather than in the generation phase. The third problem is that it is often impossible to satisfy all of the preferences at once - usually some preference rules have to be broken; but traditional descriptions of the rules do not assign the preference rules with a clear order of importance. In fact it is not at all clear that simple priorities can be given in general.

2.3 How to harmonise a chorale melody (and why it is difficult)

Typically, beginners are given a suggested order in which to carry out the tasks involved in harmonisation. The order and identification of tasks given below is adapted from Thomas (1985);

- start by deciding where to put cadences,
- write chords for the cadences,
- decide on a harmonic rhythm,
- work backwards considering all possible chords,
- choose a chord sequence from the sets of possible chords,
- work backwards fitting in the bass,
  (this must fit the chords and move properly with the melody line),
- fit the inner voices.

As more and more of the harmonisation is completed, more and more mutual constraints arise between what is being written and what has already been written. Whereas in the early stages many choices would have been more or less equally acceptable, as the task progresses, it typically gets harder and harder to fit the requirements at all, and sooner or later some set of requirements are over-constrained and cannot be met, necessitating the changing of an earlier decision. However, the earlier decision may itself have been the only choice possible at that stage, so that earlier choices have to be undone and so on. So, for example, it may be impossible for an inner voice to complete the planned chord and stay in the proper

relationship with the bass, making it necessary to change the bass, which may be impossible without changing the chord, which may necessitate changing the chords sequence and so on, in backtracking fashion.

2.4 Pedagogic use of Vivace

Given that the textbook rules for carrying out the task have been known for several hundred years, it is not, as Thomas (1985) points out, particularly difficult to write a rule-based system that implements the text book rules. Equally, it should be reasonably straightforward to construct an traditional intelligent tutoring system to use text book rules to criticise students' work and serve as a model of the expertise they are supposed to acquire. But how useful would such a tutor be, given the problems inherent in the task of harmonisation we have already outlined? Thomas not only constructed such a tutor, but also devised ways of making it pedagogically useful, despite the limitations of the text-book theory. Indeed, much of the pedagogical leverage was obtained by using the tutor to *illuminate* the limitations of the theory, as we will now consider.

By building and experimenting with Vivace, Thomas was able to establish quite firmly that typical text book rules are an inadequate characterisation of expert performance of the task. For example, Thomas discovered that if tenor and alto parts are written using only conventional rules about range and movement, the tenor voice would often move to the top of its range and stay there - which is musically ridiculous.[1] Thomas posited that there must be a set of missing rules and meta-rules to fill the gaps, and set about trying to find the rules using Vivace as an experimental tool. Note, of course, that a human musician has a role to play in each experiment, deciding on the basis of intuitions whether a result is musically acceptable or not. Thomas was able to make explicit many new detailed considerations about the task that were previously only tacit intuitions, and found out that many of the traditional rules were overstated or needed redefining. As well

[1] This echoes Rothgeb's experience in implementing the textbook rules for realising unfigured continuo bass. Rothgeb (1980) implemented complete computational models of Heinichen's and Saint-Lambert's eighteenth century procedures for bass harmonisation. The computational model, not surprisingly, showed the bodies of rules to be incomplete and inadequate.

as leading to new knowledge in matters of detail, Thomas was able to make explicit knowledge about the task at a more strategic level. For example, Thomas found that early versions of Vivace would sometimes 'get stuck', due to the preponderance of rules that prohibited actions as opposed to proposing actions. This led Thomas and her human pupils to focus on and formulate a number of heuristics for 'what to do' as opposed to 'what not to do'. Experience with Vivace also underlined for Thomas the need to make human pupils aware of high level phrase structure, before diving into detail of chord writing. Having experimentally discovered new explicit knowledge about the task, as a result of "teaching" her expert system, Thomas used this knowledge as a basis for writing a new primer for the task. Part of this knowledge was also used in a commercial program, MacVoice, which criticises students' voice-leading in a restricted version of a similar task. We will consider MacVoice after summarising the contribution of Vivace.

2.5 Summary of the contribution of Vivace

Vivace made a significant contribution to music education by enabling the received wisdom about a particular task to be procedurally run, debugged and refined, leading to better explicit knowledge of what was being taught. As well as contributing to music education, this work may be in the vanguard of a new wave of computational musicology, allowing theories of particular areas of music to be experimentally tested.

3 MacVoice: a critic for voice-leading

Macvoice is a Macintosh program arising from the expert system Vivace that criticizes voice-leading for four-part harmonisation. In order to explain what voice-leading is, we need first of all to analyse the task of harmonisation conceptually (as opposed to analysing it in terms of successive steps, as we did earlier). Readers who are familiar with the notion of voice-leading may wish to skip onto sub-section 3.1.

Four-part harmonisation has many degrees of freedom, but must be carried out within a series of constraints. One degree of freedom arises from the harmonic

ambiguity of any melody. There are many different harmonic sequences that could fit a given melody. Hence one implicit subtask in harmonisation is to choose a series of chords - although this need not be done explicitly, and may emerge from the combined activity of the individual voices. A second implicit subtask is to construct a bass line that fits the chord sequence (though the bass line may sometimes dictate the chord sequence): at the same time, the bass line must make a "good melody" in its own right. (Note that no clear, detailed guidelines for a "good" melody are known.) A third task is to write the inner voices. Two overall areas of constraint apply at all points to all voices, and it is here that the task of voice-leading arises. The first area consists of constraints on "chord construction" (Forte, 1979) i.e. preferences and prohibitions governing the pitch spacing of simultaneously sounded notes. The second area consists of constraints on "chord connection" i.e. the preferences and prohibitions on the way individual voices move with respect to themselves and other voices as they progress from chord to chord. One purpose of the chord construction constraints is to try to ensure a reasonable texture at all points - not too cluttered or top heavy. Another purpose of the chord connection constraints is to try to ensure that every voice is perceived to move independently at all times - not seemingly in a subservient relationship to some other voice. Other purposes of the constraints are to ensure that each chord is well-defined, and yet others may refer to a particular historical style. It is these two overall areas of constraint, chord construction and chord connection - particularly the constraints on chord connections - that are known as *'voice-leading'*. (Note that voice-leading is not applicable everywhere in music - it may be appropriate for a musician to ignore voice-leading constraints in some contexts, either because voice independence is not a goal at that point, or for stylistic reasons.) The purpose of MacVoice is to criticise solely this aspect of students' harmonisations.

3.1 Use of MacVoice

The MacVoice interface includes a simple music editor. It is possible to input notes in any order, for example a chord at a time, or a voice at a time or in any fragmentary fashion. As soon as any note is placed on the stave, MacVoice displays its guess as to the function of the corresponding chord in the form of an annotated Roman numeral. Two important limitations of MacVoice are as follows: firstly, all notes

must be of the same duration (so that the chords form homophonic blocks); and secondly, the piece must be in a single key. Apart from saving and loading files or erasing the stave, there is only one other menu function, called "voice-leading" (fig 25). When this is selected, a rule base of voice-leading rules inspects the harmonisation, and then provides a list of voice-leading errors. MacVoice is capable of quite flexible use, since it can be used on exercises where any combination of voices or notes are already filled in.

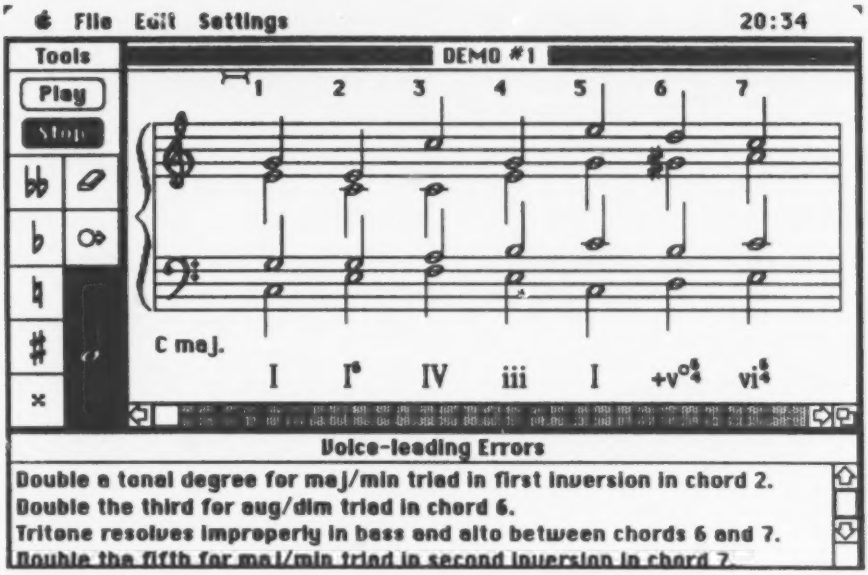


Fig 25 MacVoice

3.2 Criticisms of MacVoice and possibilities for further work

MacVoice is the first program of its kind, and is in practical use at Carnegie Mellon University. MacVoice opens up many possibilities for further work. The comments below mainly concern such possibilities. At present (version 2.0), MacVoice only points out errors, it doesn't give strategic positive advice. On the basis of Thomas's own work on Vivace, a future version could perhaps give positive advice. Neither does MacVoice address the sensibleness or otherwise of the chord sequences involved - but such a facility could in principle be added in various ways if desired.

An interesting topic for further research might be to attempt to show *visually* what the voice-leading constraints are, or what some of the preferred possibilities are at any point. This might be very difficult, if only because there are many voice-leading rules applicable at any given time. Perhaps as the task became more constrained, candidate notes might be shown (on request) in different shades of grey corresponding to more or less likely possibilities. Another possibility for further research would be to try to abstract or group the rules according to a smaller number of abstract principles that they appear to serve, and to try to use the principles to inform and illuminate the tutor's criticisms. A closely related topic for further work would be to design a tutor or environment that could help motivate for the student the origin and scope of the rules. Examples of efforts to separate lower level knowledge from higher level knowledge in expert systems, and to provide graphic windows on an existing tutor's inference processes can be found in work on Neomycin (Clancey, 1983) and Guidon-Watch (Richer and Clancey, 1985).

Another intriguing possibility for further research in ITS for voice-leading arises from recent work in the cognitive psychology of music (Wright, 1986), (McAdams and Bregman, 1985), (Cross, Howell and West, 1988), (Huron, 1988). This research seems to suggest that voice-leading rules may be a special case of the rules that govern conditions under which the human auditory system perceives distinct noises as emanating from a single entity or not. This is an important task from an evolutionary point of view for a human in a noisy forest trying to distinguish leaves moving in the wind from a hungry predator. (This image is due to McAdams (1987).) The function of the rules of voice-leading in four-part harmony may be largely to make sure that each voice is perceived as an independent entity (while each voice contributes to the purposeful chord sequence that emerges from the independent activity). It is not clear how far such general principles might go to precisely constrain traditional voice-leading rules. Such principles might have far-reaching educational applications in future intelligent tutors for voice-leading.

This completes our comments on Vivace and MacVoice. The other intelligent tutor for music composition that we shall consider is Lasso (Newcomb, 1985).

4 Lasso: an intelligent tutoring system for 16th century counterpoint

Lasso is an intelligent tutoring system for 16th century counterpoint, limited to two voices. The task tackled in writing counterpoint is similar to the voice-leading task discussed above, but with a different, historically earlier, set of stylistic constraints added. Rules for this paradigm of composition were formalised by Fux (1725) from the practice of earlier composers such as Palestrina [2]. A problem highlighted by Newcomb (1985) in teaching counterpoint, and indeed other historical styles of music, is confusion about whether the goal is to encourage good composition, or to encourage scholarly fidelity to a style as illustrated by historical documents. The goal behind Newcomb's approach to formalising the style, as incorporated in Lasso, is neither of these. Newcomb, with commendable honesty, notes that his rules are intended to serve as simple and consistent guidelines to help students know what is required to pass exams. Like Thomas, Newcomb found that the process of codification of the necessary knowledge required going beyond rules and guidelines given in text books. Unlike Thomas, Newcomb goes about this in a somewhat behaviourist, probabilistic manner, analysing scores to find out such 'facts' as "the allowable ratio of skip to non-skip melodic intervals" and "how many eighth note passages can be expected to be found in a piece of a given length" (Newcomb, 1985, page 53).

Lasso was implemented in a CAI PLATO environment and is more reminiscent of CAI than ITS. It does not appear that Lasso can write counterpoint itself (Newcomb, 1985, page 60), and Newcomb acknowledges that the knowledge used for criticising the students work is not coded particularly explicitly. It appears that each set of related criteria or rules for judging students' work is implemented as one lump of branching procedural code. Each subtype recognised by such "judging engines" is associated with unvarying canned error messages, help messages, or congratulatory messages, etcetera. The programming style of Lasso is highly constrained by the limitations of the PLATO environment. As we have already

[2] Counterpoint is a style of part writing where each voice has more or less as much melodic interest as any other voice - a more demanding task than ordinary four part harmony where the inner voices are relatively subservient to the outer voices. Exercises in counterpoint drawing on Fux's rules formed an important part of the basic education of the classical composers, and still form the basis of the way counterpoint is taught.

indicated, where the rules go beyond traditional rules, they are based on a way of characterising style that relies heavily on counting and legislating about the frequency of particular low level features.

From an ITS perspective, although its CAI flavour limits its extensibility, Lasso is nonetheless an impressive achievement. Lasso has a high quality music editor associated with it; tackles a complex musical paradigm; and has been used in real teaching contexts. Many of the limitations of Lasso, such as the inexplicitness of the knowledge representation, stem from the limitations of the PLATO environment (Newcomb, 1985). In practical terms, Lasso could give students expert criticism when access to expert criticism was otherwise unavailable (as it may be to those working alone) or limited (as it often is in large classes). This could be of substantial practical value, bearing in mind that there is a large space of 'right' answers which cannot realistically be codified in advance, and which can only be verified or criticised by an expert.

On the other hand, even if we were to imagine a reimplementation of Lasso with the knowledge coded more explicitly, able to write counterpoint itself, extended to cope with more features of the style, with a good response time and so forth, there would still be substantial problems to solve. Firstly, the rules are at a very low level, and there are a lot of them. This is reflected by the fact that there needs to be a system rule preventing more than one hundred and twenty seven comments being made about any given attempt to complete an exercise! To give a flavour of the rules, typical remarks by Lasso include;

- "A melodic interval of a third is followed by stepwise motion in the same direction."
- "Accented quarter passing note? The dissonant quarter note is not preceded by a descending step." (Newcomb, 1985, Page 58)

The corresponding low level rules are accompanied by canned texts that attempt to put the rules in a more general context and motivate them, but a user could easily be continually overwhelmed by the quantity of relevant help text required to put in

context a myriad of low-level criticisms. Students complained that "it was so difficult to satisfy LASSO's demands that they were forced to revise the same exercises repeatedly" (Newcomb, 1985, Page 60). As we suggested in the case of MacVoice, one way of tackling this problem might be try to code explicit general principles governing the low-level rules. Such codified principles might be used to cut down the number of low-level comments in any particular case and replace them by a smaller number of relevant but more general observations. (The need to represent strategic domain skills in expert systems in this way was originally exemplified by Clancey and Letsinger's (1981) work on Neomycin.) As pointed out in the previous subsection, research in the cognitive psychology of music might be a good starting point for efforts to codify the relevant abstract knowledge in this particular domain. Scope for other future research on Lasso, making use of traditional ITS techniques, might include the provision of explicit teaching rules and an explicit user model. Such developments might help the tutor to reason about when not to say things; decide to concentrate on one fault at a time in a principled way; or to reason explicitly about what strategic advice to offer on setting about the task.

5 Conclusions on intelligent tutors for music composition

Both MacVoice and Lasso are pioneering musical ITS systems[3]. They have a prescriptive and critical teaching style, and it is reasonable within their restricted domains that they should use such an approach. The traditional approach works in these cases because 16th Century Counterpoint and four-part harmonisation are genres that (as taught in schools) can be considered to have known rules, goals, constraints and bugs that are stable. A prescriptive, critical style would be far less appropriate in a tutor intended to cover a wide range of evolving styles. In such a case, it might be possible to write expert components for a *series* of sub-genres but

[3] There are not currently many intelligent tutors for music composition, and information about those that do exist tends to be sparse. Other work in the area is noted below. 'THE MUSES' is an intelligent tutor for music theory developed by Sorisio (1987). 'Piano Tutor' is an intelligent tutor for playing the piano under development by Sanchez, Joseph, Dannenberg, Miller, Capell and Joseph (1987). Baker (1988) has partially designed a tutor for aspects of musical interpretation based on work by Lerdahl and Jackendoff (1983) and Friberg and Sundberg (1986). Recent related work has been carried out by Tobias (1988); Cope (1988); Fugere, Tremblay and Geleyn (1988); and Ames and Domino (1988).

this would not satisfactorily address the problem of evolving styles. Any new piece of music worth its salt invents new rules, goals and constraints specifically for that piece, and may consequently violate some rules normally expected of the genre.

A second problem with the 'traditional' intelligent tutoring system approach is that enforcing rules abstracted from existing practice is not necessarily the best way of teaching music composition. Verbal definitions of a musical concept are often very impoverished compared with the rich multiple meanings they come to have for an experienced musician. It is all very well to define a dominant seventh, for example, (example due to Levitt (1981)) to a novice in terms of its interval pattern and then give some rules for its use, but to an experienced musician, the 'meaning' of a dominant seventh would vary greatly depending on the context - for example it could imply a forthcoming cadence, evidence of a modulation, an unexpected direction taken by a chord sequence, etcetera. Getting the novice to obey any set of rules is really far less important than making the novice aware of, and able to manipula e intelligently, the structures and expectations that are available to the more experienced musician. What is needed is not so much verbal explanations of what rules have been violated, as some way of providing structured sequences of experiences that make the novice aware of musical structures, able to manipulate them in musically intelligent ways, and capable of formulating sensible musical goals to pursue.

This completes our review of intelligent tutors for music composition. In part III of the thesis, we will propose a different kind of framework for intelligent tutors in creative domains, as a step towards solving some of the problems we have discussed.

6 Related work in other fields

In this section we give pointers to work in other fields that is related to the use of computers in music education.

6.1 Artificial Intelligence and education

A good early collection of papers on AI and education research is Sleeman and Brown (1982). Good discussions of most of the issues underlying the field can be found in O'Shea and Self (1983). A useful critical survey can be found in Elsom Cook (1984). Good recent collections of papers can be found in Self (1988a), Elsom-Cook (in press), and elsewhere. The most comprehensive up-to-date survey is Wenger (1988).

6.2 Computer musical instruments and music editors

Computer musical instruments, computer-based sequencers and computer sound synthesis are essential for serving the input/output and storage needs of aspects of the work in the thesis, but need not concern us directly. Music editors are very relevant to educational needs, but usually at a low level, as the musical equivalents of 'word processors'. At a higher level, good music editors merge with 'musician machine interfaces', already discussed. Good surveys of sound synthesis techniques can be found in Roads and Strawn (1985) and Loy and Abbott (1985). An up-to-date survey of an important class of music editors can be found in Scholz (1989). Updates in these rapidly expanding fields can be found annually in the proceedings of the International Computer Music Conference, and quarterly in the excellent Computer Music Journal.

6.3 "Computers in music research"

Confusingly, because textual analysis of scores using simple "dumb pattern matching", string search and statistical methods was among the first uses of computers in music research, researchers doing this kind of research sometimes claim that the term "computers in music research" applies exclusively to this area. Consideration of first principles in artificial intelligence and cognitive psychology make it unlikely than much knowledge of any depth will be learned about music in this way. Even practitioners (Gross, 1984) acknowledge doubts about the quality of information gained by this approach.

6.4 Computer-assisted music analysis

See Smoliar (1980) for early information on computer-assisted music analysis

(though the field has since been much influenced since by Lerdahl and Jackendoff (1983)). The complexity, subtlety and lack of precisely defined knowledge required for most kinds of music analysis tends to limit its relevance to the needs of untrained novices.

6.5 The cognitive psychology of music

Good general surveys of the *cognitive psychology of music* can be found in Sloboda (1985), Howell, Cross and West (1985) Deutch (1982), Dowling (1986). Recent research can be found in the excellent journal 'Music Perception'. The single most influential work in the field is probably Lerdahl and Jackendoff (1983).

6.6 Knowledge representation in music

Good general surveys of knowledge representation in music can be found in Roads (1984) and the regular series in Computer Music Journal ,"Machine Tongues".

6.7 Computer-aided composition

Computer-aided composition is normally taken to refer to automated assistants or environments for practiced composers who already have a lot of tacit knowledge about music, and well-developed compositional skills and strategies. It also frequently refers to composition in the electro-acoustic paradigm. Experienced composers are notoriously diverse in their approaches (Sloboda, 1985). Hence much of this work concentrates on providing powerful tools that make as few constraining assumptions about musical structure as possible.

7 Conclusions on the use of computers in teaching music.

To complete part I, we will review our conclusions in three areas, namely: early approaches to the use of computers in teaching music; developments in the musician-machine interface; and intelligent tutors for music composition. These correspond to collected and more concise versions of conclusions already drawn in the relevant sections.

7.1 Conclusions about early approaches

Computers are currently used in music education mostly as quiz-masters, easily used musical instruments, "musical typewriters" and for musical games; they can they can count and identify simple musical events, and perform simple musical manipulations such as reordering, transposition, and pitch and time transformations.

High quality exploratory microworlds for music are beginning to introduce opportunities for exploratory learning for those without traditional musical skills.

Few of the early applications exploit new opportunities to make *explicit* representations of the kinds of musical knowledge that a novice may need when learning to compose tonal music. Musical knowledge is mostly left implicit.

A human teacher with expert musical knowledge could use musical microworlds to devise individually relevant experiments to help students gain the experience they need. The teacher could also help interpret the results of the experiments to students. To some extent, the same purpose might be served by books of experiments associated with particular microworlds, but this would miss opportunities to capitalise on individual differences. Intelligent tutors to complement microworlds could help with such problems.

7.2 Conclusions about developments in the musician-machine interface

In the case of experienced composers, a good design consideration is to design interfaces that have maximum flexibility and make as few assumptions as possible about how a composer will work. This approach is consonant with the observed diversity of the way in which composers work (Sloboda, 1985). However, if we wish to educate novices in some particular aspect of music, it may be useful to reverse this design consideration and build into interfaces very specific metaphors for particular aspects of music.

New developments in the musician-machine interface are beginning to provide

novices with diverse power tools for music composition, some of which we summarise below. *Music editors* make it possible for novices without traditional musical skills to manipulate and listen to music represented in common music notation. *Graphical programming kits for music,* although currently in the experimental stage, offer tools for novices to construct computational models of musical processes. *Interactive composition tools* make it easy for novices to control musical manipulations such as reordering, transposition, and pitch and time transformations in real-time. However, these exciting developments leave many issues unaddressed, which can be summarised as follows. *Limitations of Music editors:* in the case of novice musicians who want to learn to compose, the use of common music notation as a sole or chief notation may not always be appropriate, since much important information is only represented implicitly in the notation, and novices can have very great difficulty learning it. To exploit new opportunities to help beginners get initial musical experience, there is a need for other kinds of notations, and other kinds of interfaces. *Limitations of graphical programming kits for music:* visually constructed programs can be very hard to understand, unless they are very small, or unless good visual structuring facilities are provided. Graphical programming kits provide tools for building explicit computational models of musical processes, but they do not in themselves constitute such models. To use graphical programming kits to represent specific musical knowledge perspicaciously, they require augmentation in the following areas;

- identification of appropriate musical knowledge for use in models,
- formulation of the knowledge in appropriate computational form,
- selection of grain size for the knowledge to aid perspicacity,
- identification of good computational metaphors to make the knowledge easy to communicate to novices,
- devising of mechanisms for making the knowledge flexibly accessible to students from different viewpoints.

Several of these issues will be addressed in part III of the thesis. *Limitations to interactive composition tools:* although novices can undoubtedly produce interesting and varied music with interactive composition programs, there is little

evidence that their newly found "skills" would make them any more able to compose when deprived of the program, or able to articulate in any detail what they had been doing. With careful selection of initial materials, structured guidance, and expert interpretation, novices might be led to experience beneficial educational "shocks" (such as in more traditional microworlds) that would lead to the development of compositional insight. There is a role for intelligent tutors to complement such tools.

7.3 Conclusions about intelligent tutors for music composition

The traditional remedial ITS approach can be applied to certain restricted areas of music which can be considered to have identifiable and stable rules, goals, constraints or bugs. A prescriptive, critical style is less appropriate in a tutor intended to cover a wide range of evolving current practices. Any new piece of music worth its salt invents new rules, goals and constraints specifically for that piece, and may consequently violate some rules normally expected of the genre. In such a constantly evolving domain, it is hard to see how any intelligent tutoring system could know enough to be prescriptive or critical without the risk of crushing innovative and interesting ideas.

Verbal definitions of a musical concept are often very impoverished compared with the rich multiple meanings they come to have for an experienced musician. Getting the novice to obey the rules is may be far less important than making the novice aware of, and able to manipulate intelligently, the structures and expectations that are available to the more experienced musician. What is needed is not so much verbal explanations of what rules have been violated, as some way of providing structured sequences of experiences that make the novice aware of musical structures, able to manipulate them with musical intelligence, and capable of formulating sensible musical goals to pursue.

PART II: Harmony Space: an interface for exploring tonal Harmony

"I played piano at school, but not for very long. The teachers always wanted you to play "Ugly Duckling" over and over again... it drove me up the wall. What I really wanted was someone to tell me why some notes sounded better than others did...Rock'n'Roll theory, if there is such a thing".

Seamus Behan (1986), member of rock group "Madness".

"[The property of uneven spacing in a scale]... enables the listener to have, at every moment, a clear sense of where the music is with respect to such a framework. Only with respect to such a framework can there be things such as motion or rest, tension and resolution, or in short the underlying dynamisms of tonal music. By contrast, the complete symmetry and regularity of the chromatic and whole tone scales means that every note has the same status as every other. The fact that for such scales there can be no clear sense of location, and hence of motion is, I believe, the reason that such scales have never enjoyed wide or sustained popularity as a basis for music."

Shephard (1982) quoted in Sloboda (1985, page 255).

In Part II of the thesis, we present a theoretical basis and a detailed design for a computer-based interface, "Harmony Space", designed to enable novices to compose chord sequences and learn about tonal harmony. The interface is based on recent cognitive theories of how people perceive and process tonal harmony. Two different versions of Harmony Space, based on two different theories, are presented in chapters 5 and 7.

Chapter 5: Harmony Space

In this chapter, we present a version of Harmony Space based on Longuet-Higgins' (1962) theory of tonal harmony. We investigate the representation of chords, key areas, tonal centres and modulation in a simplified version of the theory. We then outline the design of Harmony Space, and investigate the representation of various chord progressions of fundamental importance using Harmony Space. The structure of this chapter is as follows;

1 Longuet-Higgins' theory of harmony
2 The 'statics' of harmony: chord and key structure
  2.1 Keys and modulation
  2.2 Chords and tonal centres
3 A computer-based learning environment
  3.1 The minimal features of Harmony Space
  3.2 Harmony Space design decisions
4 The 'dynamics' of harmony: representing harmonic succession
  4.1 One chord harmony
  4.2 Two chord progressions
  4.3 Three chord progressions
  4.4 Harmonic trajectories and paths
    4.4.1 Cycles of fifth trajectories
    4.4.2 Scalic trajectories
    4.4.3 Chromatic trajectories
    4.4.5 Progressions in thirds
  4.5 General mapping of intervals in 12-fold space
  4.6 Summary of 'dynamics' of harmony
  4.7 Harmonic plans
5 Analysing music informally in Harmony Space
6 Conclusions

1 Longuet-Higgins' theory of harmony

Longuet-Higgins' theory of harmony (1962) investigates the properties of an array of notes arranged in ascending perfect fifths on one axis and major thirds on the

other axis (fig. 1). This representation turns out to be a good framework for theories explaining how people perceive and process tonal harmony (Steedman, 1972, Sloboda, 1985). Our focus here is on the rather different goal of applying the theory to develop new educational tools. Longuet-Higgins' (1962) theory asserts that the set of intervals that occur in Western tonal music are those between notes whose frequencies are in ratios expressible as the product of the three prime factors 2, 3, and 5 and no others. Given this premise, it follows from the fundamental theorem of arithmetic that the set of three intervals consisting of the octave, the perfect fifth and the major third is the *only* co-ordinate space that can provide a unique co-ordinate for each interval in musical use (Steedman, 1972). We can represent this graphically by laying out notes in a three dimensional grid with notes ascending in octaves, perfect thirds and major fifths along the three axes[1]. We can think of columns of notes in more and more distant octaves stacked on top of each note in Fig 1. The octave dimension is discarded in most discussions on grounds of practical convenience for focussing on the other two dimensions, and on grounds of octave equivalence (Fig 1). 'Octave equivalence' simply refers to the fact that for many elementary musical purposes[2], notes an octave apart are considered by musicians to be equivalent.

2 The 'statics' of harmony

[1] In Longuet-Higgins' presentations of the theory, and in all discussions of it in the psychological literature, the convention is that ascending perfect fifths appear on the x-axis and the ascending major thirds on the y-axis. For educational use, we reverse this convention on three grounds. Firstly, it could allow students to switch more easily between versions of Harmony Space using Balzano's representation (chapter 7) and the 12-note version of the Longuet-Higgins representation. (Under our convention, the x-axes of the two representations coincide and the y-axes are related as if by a shear operation.) Secondly, it makes the dominant and subdominant areas coincide with Schoenberg's (1954) dominant and subdominant regions (though of course the x-axes and the overall meaning of the respective diagrams differ). Thirdly, the V-I movements that dominate Western tonal harmony at many different levels become aligned with physical gravity in a metaphor that may be useful to novices. This convention has since been adopted by the innovative music educator Conrad Cork(1988), to whom I introduced Longuet-Higgins' theory. Cork is the only other person I am aware of who has used Longuet-Higgins' theory educationally.

[2] As Henkjan Honing (1989) has pointed out, it can be useful to distinguish between melodic and harmonic octave equivalence. For the purposes of our discussion, we consider inverted harmonic functions to be equivalent. Later, we will outline how inversions can be controlled and represented naturally in Harmony Space.

2.1 Keys and modulation

We begin by looking at how various 'static' relationships in harmony appear in this representation. In diagrams such as Fig. 1, all of the notes of the diatonic scale are 'clumped' into a compact region. For example, all of the notes of C major, and no other notes are contained in the box or window in Fig. 1.

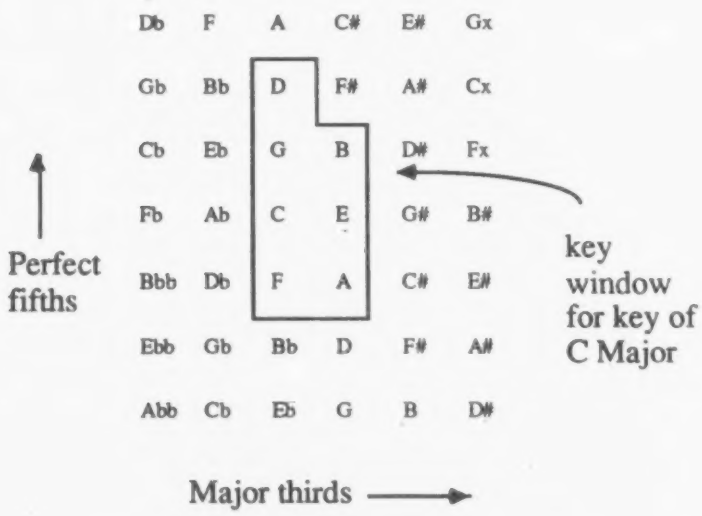


Fig 1 Longuet-Higgins' note array
(diagram adapted from Longuet-Higgins (1962))

If we imagine the box or window as being free to slide around over the fixed grid of notes and delimit the set of notes it lies over at any one time, we will see that moving the window vertically upwards or downwards corresponds to modulation to the dominant and subdominant keys respectively. Other keys can be found by sliding the window in other directions. The tonic and other degrees of the scale correspond to fixed positions in relation to the window. Despite the repetition of note names, it is important to note that in the original theory, notes with the same name in different positions are not the same note, but notes with the same name in different key relationships (Steedman (1972) calls such notes "homonyms").

However, for the purposes of educating novices in the elementary facts of tonal harmony, it turns out to be convenient to map Longuet-Higgins' space onto the twelve note vocabulary of a fixed-tuning instrument, resulting in a 12-note, two-dimensional version of Longuet-Higgin's representation. (We will look at the

reasons behind this move in Chapter 6. For the purposes of this chapter, we will just investigate its consequences.) As a consequence of this mapping, we lose the double sharps and double flats of fig 1, and the space now repeats exactly in all directions (fig 2).

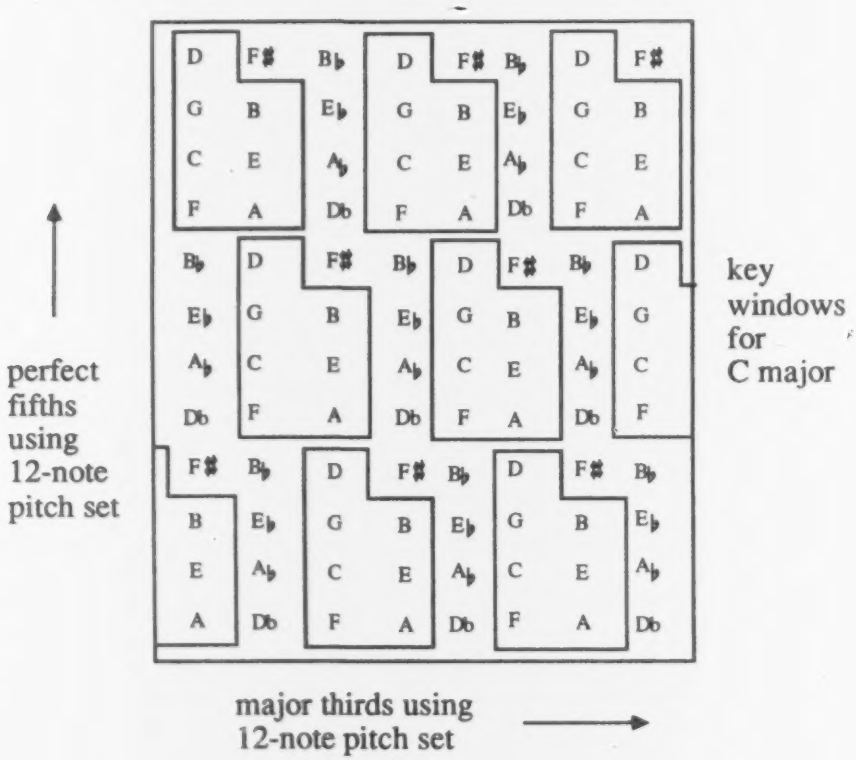


Fig 2 12-fold version of Longuet-Higgins' note array

Notes with the same name really are the same note in this space. A little thought will show that the space is in fact a torus, which we have unfolded and repeated like a wallpaper pattern. The key window pattern has been repeated as well (fig 2). (We have used arbitrary spellings in these diagrams (e.g. F# instead of Gb etcetera), but we could equally easily use neutral semitone numbers or any preferred convention.) We will refer to the original space as *"Longuet-Higgins' original space"* or the *"non-repeating space"*, and to the 12-fold space variously as the *"12-fold"*, *"12-note" or "collapsed"* space.

The collapse to the 12-fold space makes it apparently impossible to make distinctions about note spelling that could be made in the original space [3], but we can console

ourselves with the thought that in this respect it is no more misleading than a piano keyboard. We will see in chapter 7 that in many other respects it represents harmonic relationships very much more explicitly than a piano. We would encourage a novice learning about music to play with a piano to her heart's content - no one would dream of discouraging this on the grounds that it misrepresented note spelling or other harmonic distinctions.

2.2 Chords and tonal centres

Let us now turn to look at the representation of triads and tonal centres in the 12-fold space. The representation allows many interrelationships between chords and tonal areas to be summarised graphically and very succinctly, as we will now demonstrate. In 12-note versions of Longuet-Higgins' space, major triads correspond to triangle-shapes (fig 3). Major triads consist of three distinct objects in the space that are maximally close[4]. The dominant and subdominant triads are seen to be triads maximally close to the tonic triad. It is plain from the diagram that the three primary triads together contain all of the notes in the diatonic scale. There is a clear spatial metaphor for the centrality of the tonic - the tonic triad is literally the central one of the three major triads of any major key[5]. We can make similar observations for the minor triads. Minor triads correspond to rotated triangle-shapes. Like major triads, they are maximally compact three-element objects[6]. The three secondary triads generate the natural minor (and major) scale (We could deal with the harmonic minor and melodic minor scales by introducing variant key window shapes, but we will not pursue this here). The space also gives a clear visual metaphor for the centrality of the relative minor triad among the secondary triads. The 'centrality of the tonic' argument as applied to the tonal centre of the minor mode is borrowed from Balzano (1980). (We will see in chapter 6 that it is not valid in the original Longuet-Higgins' space, but applies in the 12-note Longuet-Higgins version.) Completing the full set

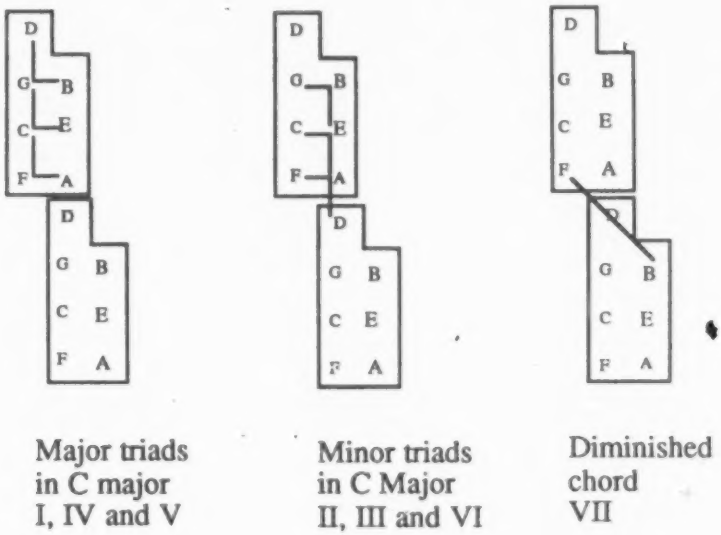
[3] But see chapter 6 for a way of recovering these distinctions.

[4] Closeness is measured by the sum of the geometric distance between note-points.

[5] Note that this argument does not depend at all on how the key windows in the 12-fold space are 'sliced up' - although this may help emphasise the relationship visually. The tonic is spatially and intervalically central in *any* grouping of three major triads that are arranged in perfect fifths in any major key.

[6] But not all maximally compact three element objects in this space are triads. For example, the constellations CFE and CEB are compact but are not triads.

of scale-tone triads for the major scale, the diminished triad is a sloping straight line (two major thirds 'on top' of each other). Note that there is a simple, uniform correspondence between shape and chord quality for triads (as represented in their maximally compact form in the 12-fold space) for any triad, in any key. (We will see in chapter 6 that this is not the case in the original Longuet-Higgins' representation.)



Major triads in C major I, IV and V

Minor triads in C Major II, III and VI

Diminished chord VII

Fig 3 Triads in the 12-fold space

Seventh and ninth chords (which are used consistently in place of triads in some musical dialects - e.g. jazz) have similarly memorable and consistent shapes in the space (Fig 4).

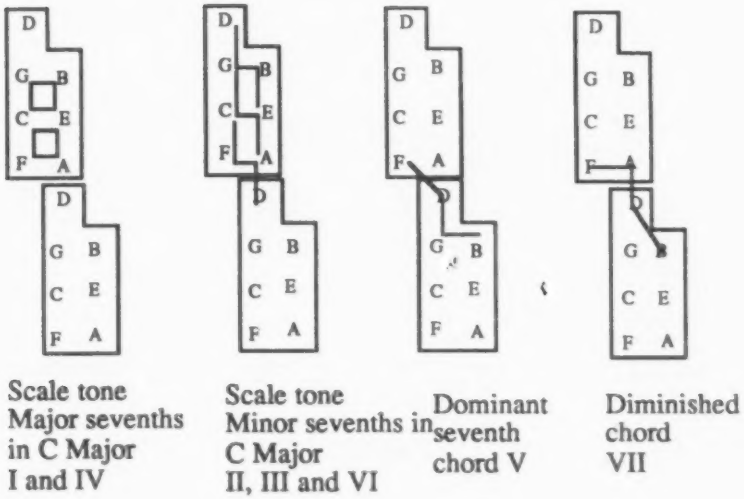


Fig 4 Scale tone sevenths in 12 fold space

3 A computer-based learning environment

3.1 A description of the minimal features of Harmony Space

We will now present the general design of a computer-based learning environment based on the theory as described so far. In chapter 7, we will discuss a restricted implementation of this design. A grid of circles is displayed on a computer screen, each circle representing a note. The notes are arranged as in Fig 2. There are two pointing devices such as mice (or one mouse and a set of arrow keys) and optionally some foot-operated switches or pointing devices connected to the computer. A mouse controls the location of a cursor. Provided the mouse button is down at the time, any note-circle the cursor passes over is audibly sounded. When the mouse button is released or leaves a note circle, the note is audibly released. More generally, we can set the mouse to control the location of the root of a diad, triad, seventh or ninth chord. (We will call the number of notes in the chord its '*arity*', borrowing from mathematical usage.) The arity of chord controlled by the mouse can be varied using a foot switch (or control keys). As the root is moved around, the quality of the chord changes appropriately for the position of the root in the scale: so for example, the chord on the tonic will be a major triad (or major seventh if the arity is set to 'sevenths') and the chord on the supertonic will be a minor triad (or minor seventh in the sevenths case). We refer to these chord qualities as the *default* chord

qualities for the arity and degree of the scale. Of course, default chord qualities may sometimes need to be overridden, and this is controlled using a second pointing device and a menu of chord qualities (or a set of function keys).

Although the qualities of chords are assigned automatically (unless overridden) as the root is moved by the user, there is a clear visual metaphor for the basis of the choice; the shape of the chord appears to change to fit the physical constraint of the key window. A second mouse or pointing device is assigned to move the key window.[7] Moving this pointer corresponds to changing key. If, for example we modulate by moving the window while holding a chord root constant, the chord quality may change. Once again there is a clear visual metaphor for what is happening, since the shape of the chord will appear to be "squeezed" to fit the new position of the key window. The environment is linked to a synthesizer and everything we have described can be heard as it happens. This general description constitutes the 'core abstract design' or minimal specification for the learning environment we shall refer to as ***Harmony Space***. Let us now itemise and comment on the design decisions of the minimal specification in greater detail.

3.2 Harmony Space design decisions

In the minimal specification of Harmony Space, notes are arranged on a computer screen in ascending major thirds (semitone steps of size 4) on the x-axis, and in ascending perfect fifths (semitone steps of size 7) on the y-axis. As a pointing device (e.g. mouse) is moved around, a cursor moves on the screen. When the mouse button (or other momentary activator on a pointing device) is pressed, if the cursor is currently in a note circle, the appropriate note-circle lights up and sounds. Letting go of the button or moving the pointer out of a circle is like letting go of a synthesiser key - depending on the timbre in question, the sound stops or decays appropriately. (Ideally, the trigger control should be pressure and velocity sensitive, like some synthesiser keyboards, to allow control of volume and timbre.) Similarly, if the button is *kept* depressed and the pointing device swept around, the appropriate

[7] The pointing device could be foot-operated (foot-operated pointing devices are known as 'moles' (Pearson and Wieser, 1986)). Alternatively, the same pointing device could be used at different times to overide chord quality and to move the key-window.

succession of new notes is sounded and appropriately completed (or held if an optional *sustain pedal* is depressed).

Clearly it is useful to allow one button depression to control chords, as well as single notes. The interface is such that, depending on the position of a separate selection device controlling the number of notes in the chord (operated by the other hand or a foot slider) each button depression produces a note, diad, triad seventh or ninth chord. The pointer indicates the location of the root of the chord. However, in common chord sequences, the chord quality is continually changing. If this had to be carried out manually at each step, perhaps by selecting labelled buttons with a separate pointing device operated by the other hand, fluent use of the interface would demand considerable manual co-ordination and musical knowledge or experience that would be beyond most novices.

A very simple design decision that solves this problem involves treating the key window as a visible, constraint enforcing mechanism. We simply select whether we want (at any time, and until further notice) chords of arity one, two, three, four, five etc. (i.e single notes, diads, triads, sevenths ninths). The *quality* of the chord is then chosen automatically on the basis of the arity of the chord and the location of the pointer relative to the key window - i.e on the basis of the degree of the scale.[8] Although in these circumstances the chord quality is being chosen automatically, the constraining effects of the walls of the key window provide a visible explanation for how the qualities are being chosen. The traditional rule for constructing 'normal' chord qualities of any arity on each degree on the scale, 'stacking thirds', has an exact visual analog in the space - the chords have to fit in the key window, start on the selected note, and are built up by extending 'East' (major thirds) or 'Northwest' (minor thirds), whichever is available (fig 3). The structure of the diatonic scale is such that there is only ever one choice. As alternatives to automatic selection of chord quality, the chord quality can be overidden at any time, or made fixed.

The key window idea lends itself to another important design decision, the idea of a *second pointing device controlling the position of the key window*. This decision, in conjunction with the visible constraint enforcing mechanism, gives rise to a single,

[8] To be strictly accurate, the chord quality also depends on the mode and the 'chord quality scheme' in use. These refinements are described in detail in appendix 1.

simple, powerful and uniform spatial metaphor for the intuitive physical control and visualisation of chord sequences. The presence of two pointing devices means that, for example, we can hold a chord root steady while moving the key window, and the chord quality will change appropriately.

4 Representing harmonic succession

So far, we have used Harmony Space to look at the representation of key areas and chords. Let us now move on to investigate and illustrate harmonic succession and progression as represented (and as playable using simple gestures) in Harmony Space. Comments in italics are used to summarise key points about the expressivity of the interface. (See Appendix 3 for the chord notation convention to be used.)

4.1 Harmony with one chord

One of the simplest harmonic accompaniments possible for a song is a tonic pedal or bass drone. (Note that we are talking about explicit harmonisation of a melody rather than the harmony implicit in an unaccompanied melody). This harmony involves the repetition or holding of a single note, diad or triad as an accompaniment to a melody. This can be a very suitable accompaniment for countless songs (e.g. the French folk songs "Patapan" and "J'ai du bon tabac" (Campbell, 1985)). The repetition of the tonic chord (major or minor as appropriate) can also be used as a simplified accompaniment to harmonically more complicated pieces when starting to learn about practical harmony (Campbell, 1985, chapter 1). Many examples of extended use of a tonic pedal by serious composers can be found. Campbell (1985) cites the 8th and 22nd of Brahms' variations on a theme of Handel and part of Mendelssohn's Variations Serieuses numbers 11 and 5. *In Harmony Space, a repeated tonic accompaniment can be seen to stay in the 'home' position, at the centre of the three major (or minor) triads of the key window* (Fig 3).

4.2 Two-chord progressions

A great many pieces can be played making use of just the tonic and dominant chords. For example, the children's songs "Michael Finnigan", "Oranges and Lemons", "Skip to my Lou" and the carol "I saw three ships" (Campbell, 1985). To pick some other examples more or less at random, Aldwell and Schachter (1978, page

75) cite a sonata by Kuhnau (1660 - 1722) using only I and V. The same authors note (page 79) the frequent use of I V I in the minor in Bach chorales. If we widen our scope to include I, V and V7, (a four note version of the V chord) the harmony of a multitude of pieces and fragments can be covered. For example, Aldwell and Schachter (1978) cite part of the Haydn Symphony No. 97, III. The V I chord progression is the basic progression of western music (Aldwell and Schachter, 1978 page 75). *In Harmony Space, the harmonically very important I V I progression can be seen as one that begins on the central major triad of the key, and then moves to a maximally close neighbour before returning 'home'* (fig 3).

4.3 Three chord progressions

The next vocabulary for chord succession we shall look at, (I IV V), major or minor, is adequate for a great many whole pieces. Cannel & Marx (1976) refer to this as the 'Elementary Classical Pattern'. It is also so widespread in popular music (e.g. Chuck Berry's (1958) "Johnny B. Goode" and Taj Mahal's "Six days on the Road") that it has a well-known vernacular name - the "three chord trick". It is also the vocabulary for an entire genre of vernacular music - the 12 bar blues. *In Harmony Space, this sort of harmonic progression is seen to move either side of the tonal centre by the smallest possible step and then return 'home' to the centre.*

IV and V can provide different harmonic contexts give variety to similar melodic material. The sense of I as tonal centre is so strong that chord progressions frequently begin, and almost always end, on I unless the composer is deliberately trying for a sense of incompleteness or shock. Note that this provides a way for a novice (or any other composer) to make it obvious that a chord progression is temporarily pausing, but has not finished (by pausing on V or IV).

4.4 Harmonic trajectories

4.4.1 Cycles of fifths

We now come to a type of harmonic progression that is both widely used and of great theoretical importance in tonal music. This kind of progression involves establishing a tonal centre, jumping to some non-tonic triad and then moving back to the tonic in a straight line through all intermediate triads along the cycle of fifths axis.

We will use the term *harmonic trajectory* to refer to any straight line motion in harmony space to a tonal goal (normally the tonic). (This terminology is in deliberate imitation of Levitt's (1985) use of the term 'trajectory' to refer to the related phenomenon of melodic trajectory - i.e scalic passages (possibly decorated in various ways) towards hierarchically important notes.) Cannel and Marx (1976) label this kind of harmonic movement (expressed in different terms) as 'classical harmonic movement', and claim that it lies at the root of most popular music, as well as music in the classical and baroque periods. Depending on the length of the initial leap, some examples of this kind of progression are;

II V I

VI II V I

III VI II V I

Pieces using such progressions are ubiquitous. The chord progression of Bach's prelude No 1 in C Major (1968/~1722) begins;

I II7 V7 I VI II dom7 V I ma7

giving two harmonic trajectories in succession (The sharpened leading note in the II dom7 chord indicates that the key window may have shifted. There is an ambiguity about whether the second trajectory is aimed at the initial tonal centre I or the tonal centre of V being prepared). The chord sequence of 'I Got Rhythm' by George and Ira Gershwin (1930) is made up of repeated trajectories of the form;

VI II V I

Conrad Cork (1988) in his Lego Theory of Harmony identifies the progression I VI II V I as a standard building block of popular songs which he labels the "Plain Old Turnaround". Trajectories can be elaborated or camouflaged. Cannel and Marx (1976) give a brief taxonomy of elaborations and camouflages for this basic progression, of which three variants are; seeming to miss harmonic points in the trajectory but going back to state them, starting a trajectory without having initially stated the triad of the tonal centre, and insertion of foreign non-related chords for surprise effect. Closely related chords (e.g diminished or augmented chords) are

sometimes substituted for particular chords. As the brief Bach example above hints at, variations can be introduced by moving the tonal centre in mid-trajectory. All of these progressions (II V I, VI II V I, III VI II V I, etcetera), correspond to straight lines vertically downward in the 12-fold space with the tonal centre as their target (Fig 6a). Harmonic trajectories vertically downwards in a cycle of fifths from tonic to tonic fall into two classes, as follows. (The same applies to all sufficiently long harmonic trajectories.) In 'tonal' cycles of fifths, only notes in the current key are played. This corresponds to a straight line that bends or jumps where necessary to remain in the key window (Fig 6b, and 6c excluding shaded points). In the key of C, for example, the jump is the diminished fifth from F to B. In the case of 'real' cycles of fifths, the root moves in a straight line down the perfect fifth axis, cutting across chromatic areas outside the key window (Fig 6c including shaded points).

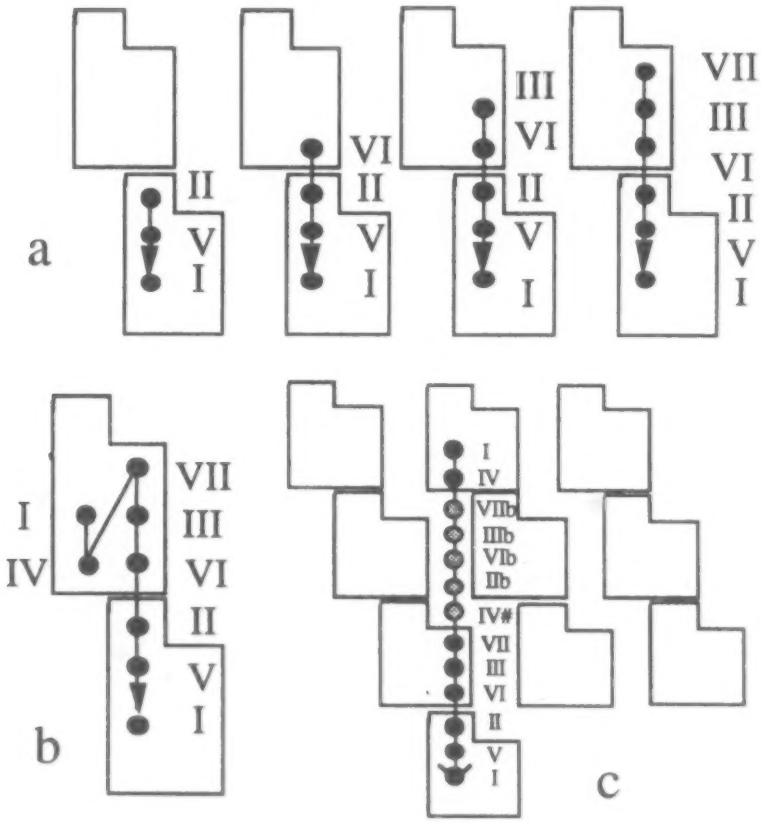


Fig 6 Cycles of fifths in 12 fold space

In Harmony Space, we can play either kind of tonal cycle of fifths by making a

single vertical straight line gesture. The chord quality can be seen and heard flexing to fit within the key window (figs 3 and 4). This works even if there are modulations (movements of the key window) in mid-chord sequence. To play a real cycle of fifths, we simply switch on the option allowing roots to be sounded in the chromatic area outside the key window. Note that for roots outside the key window, there are not necessarily any obvious default chord qualities. In some styles of jazz, these chords are more or less consistently played with dominant seventh quality, though more generally, the quality is influenced by the melody. To deal with this, either some suitable fixed chord quality must assigned to chords with chromatic roots, or chord qualities must picked manually.

If we reverse the direction of movement on the cycle of fifths axis and consider chord sequences moving vertically upwards, we have, following Steedman's (1984) terminology, *extended plagal sequences* and cadences. Plagal cadences (i.e. 'IV I' endings) are commonplace, but extended chord sequences of this sort (i.e. chord sequences like I V II VI etc.) are rare, for reasons Steedman (1984) investigates. Extended plagal chord sequences are occasionally used as the basis for entire pieces of music, for example "Hey Joe" (popular arr. Jimi Hendrix). *This kind of progression corresponds in harmony space to simple straight line gestures along the fifths axis, but in the upwards direction.*

<u>I V II VI III</u> Hey Joe (popular arr. Jimi Hendrix)

4.4.2 Scalic trajectories

Bjornberg (1985) has identified uses of modal harmony in recent popular music where scalic progressions (i.e. progressions up and down the diatonic scale) play an important role. Bjornberg's theoretical insights may be well-suited to help novices using Harmony Space understand the uses of modal materials in current popular music. Bjornberg distinguishes between the following harmonic practices;

- Standard functional harmony as used in the classical style.

- Jazz harmony (closely related to functional harmony (Mehegan,1959)).

• Unfunctional harmony, where a sequence of (typically major) triads is grafted onto a melody in a non-functional way.

• 1960's modal jazz, (e.g. "So What?", (Miles Davis , 1959) where all the chords are based on a single (or "modulating") modal scale, but functional harmony may be repudiated by the use of non-triadic chords such as combinations of fourths. The modal scale is explicitly stated and consciously used to guide melodic and harmonic elaboration.

• Contemporary modal rock harmony, as described below.

Contemporary modal rock harmony is typically triadic, with the modal scale not necessarily strongly present in the melody, but apparent from the chordal material. Modal harmony involves an attempt to shift the perceived harmonic centre from the major or minor tonal centres. Balzano (1980) has argued that the major/minor tonics tend to be 'built into' diatonic materials whenever triads are used. Consequently, to establish a modal harmonic centre, explicit counterbalancing measures must be taken This is usually achieved by devices such as: the use of pedal notes; repetition of the desired central chord or note; emphasis by prominent metrical placing; and sometimes by employing characteristic harmonic progressions that are typical of modal harmony. Described in terms of Harmony Space imagery, these characteristic progressions are typically *scalic trajectories* (fig 7) to the modal centre in question (e.g. VI VII I or III II I), or *scalic oscillations* away from and back to the centre (e.g. I VII VI VII I or I II III II I). Collectively, we can refer to these modal progressions as *modal harmonic ostinati* (Bjornberg's term) or *"mho's"* for short. (An ostinato is a short, repeated musical pattern.) Bjornberg almost exclusively considers *Aeolian* rock harmony, which he notes has become increasingly prevalent in the last ten years[9]. Examples include;

All along the watchtower (Bob Dylan, 1968), <u>i bVII bVI bVII i</u> [10]

[9] Bjornberg suggests that the goal oriented progressions of functional harmony symbolise the ideology of industrial capitalism, while the increasing use of aeolian harmonic ostinati reflects a more aimless, shifting attitude on the part of youth, and a rejection of, or ambivalence towards, goal-oriented industrial culture.

[10] We have adopted the convention that chord sequences notated using the classical convention (see

Layla (Eric Clapton, 1970),	i bVI bVII i
Fox (Hawkins, Holland, Clarke, 1980),	i ii dim bIII
Sultans of Swing (Dire Straits, 1978),	cadential bVI bVII i
Message in a bottle (Police, 1979),	aeolian I VI VII IV
In the Air Tonight (Phil Collins, 1981),	aeolian I VI VII VI I
Lets Dance (David Bowie, 1983).	(aeolian, but not an mho)

Not discussed by Bjornberg, but also becoming increasingly prevalent is the use of Dorian rock modal harmony. Recent examples include;

Michael Jackson's (1982) "Billie Jean"	i ii III ii
Keef Hartley's (1969) "Outlaw"	i ii III ii
Stevie Wonder's (1973) "Living for the City"	i ii III ii
Steely Dan's (1974) "Pretzel Logic"	i ii III ii
Soft Machine (1970) "Out Bloody Rageous"	III ii i

Like cycles of fifths, both kinds of modal harmonic ostinati, in whatever mode, correspond in Harmony Space to simple straight line gestures. However, the straight lines for modal harmonic ostinati are on a different axis to cycle of fifths progressions: in the 12-fold space, scalic sequences (i.e. movement up and down the diatonic scale) correspond to diagonal trajectories constrained to remain within the key windows (fig 7). Clearly, the brief sequences of scalic progressions that occur in traditional and jazz harmony can also be represented in the same way. (Note that scalic sequences can also be produced by moving down the cycle of fifths axis omitting every other note. This emphasises the relationship between cycles of fifths and scalic movement but is harder to play fluidly on the interface because of the jumps involved.)

appendix 3) are underlined. Chord sequences using Mehegan's (1959) convention are not.

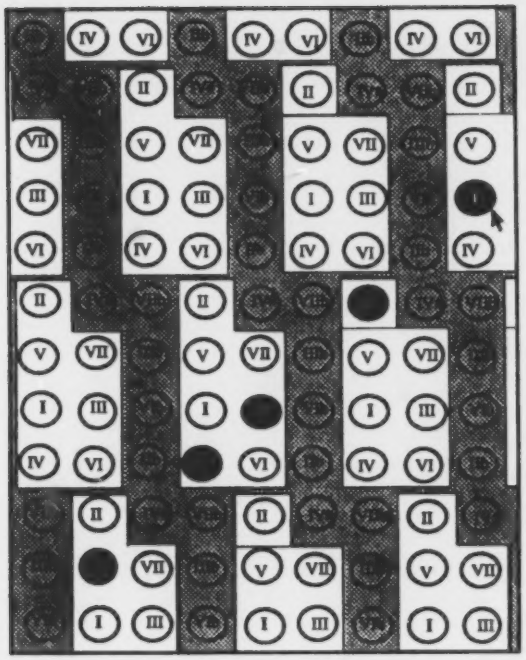


Fig 7 A scalic progression

4.4.3 Chromatic progressions

If the constraint is removed that the root must stay within the key window, scalic sequences become chromatic sequences (fig 8). Chromatic chord succession happens a lot in jazz and other vernacular styles influenced by jazz, although it would seem rather out of place in some traditional tonal contexts. Given our focus, it is an important and useful progression. Songs using this progression include, "Tea for Two" (Youmans, Harbach, Caesar, 1924), "I don't need another girl like you" (Walsh, 1976), All the things you are" (Kern-Hammerstein) and "Girl from Ipanema" (Jobim and Moraes, 1963). Many more examples can be found in Mehegan (1959). Some progressions commonly found in jazz are given below in Mehegan's notation (see appendix 3 for notes on the notational convention).

II bIIx I

III bIIIx II bIIx I

I #Io II #IIo III

III bIIIo II bIIM I

bVø IVo III bIIIo II bIIx I (Mehegan, 1959)

This kind of chromatic movement is strongly related to cycle of fifths, in that if alternate chords are ignored, the progressions coincide. For example:-

III bIIIo II bIIM I (chromatic progession)

III VI II V I (cycle of fifths)

Indeed, in jazz standards, chromatic progressions often stand in the place of cycles of fifths. Jazz players refer to the substitution of the chromatic chords for the chords in a cycle of fifths progression as "tritone substitution". This is so called because the root of the substituted chord (e.g. bII) is a tritone away from the root of the original chord (e.g. V). *In Harmony Space, chromatic progressions can be played by the same straight line gesture as diatonic progressions, if the constraint is removed that the root must stay within the key window.* Either some suitable fixed chord quality must be assigned to chords with chromatic roots, or chord qualities must be picked manually.

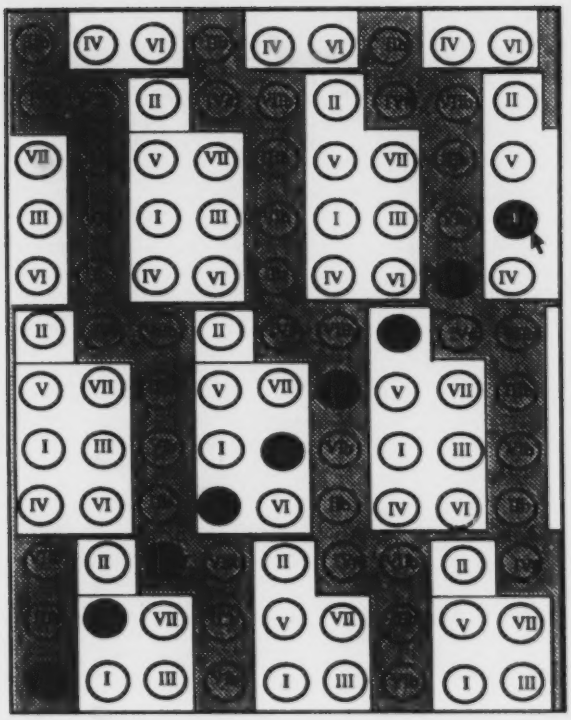


Fig 8 A chromatic progression

4.4.5 Progression in thirds

Harmonic progressions by thirds are not nearly as common as cycles of fifths in tonal music, but they are sometimes used. The most common version is "diatonic progression in thirds", which can be identified with subsequences of the sequence VI IV II VII V III I. Such progressions can be visualised in Harmony Space as shown in fig 9. Such sequences are often used in the descending direction. This pattern is the same as thc "chord growing rule" pattern going backwards. An example can be found in "Ballade" (Mehegan , 1960 Page 4), whose chord sequence is as follows;

I III II / VI / VI IV II / VII / V III I / VI /
III IV V / VI IV II / V III I / IV II III IV / II III / I / .....

"Ballade" (Mehegan 1960, Page 4).

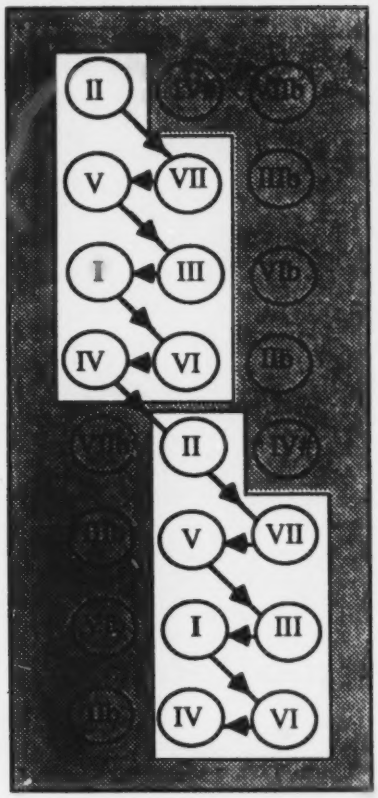


Fig 9 "diatonic cycle of thirds"

The importance of the major third interval in the definition of the 12-fold space may lead a novice to wonder whether a horizontal movement in thirds *unconstrained* by key window has an important function in tonal music (Fig. 10). An example of such a progression is I III VIb I. This kind of progression does not appear to have an important role since the progression fails to establish itself in any particular key. Similar problems occur with any attempted progression in strict minor thirds (fig 11). The pattern repeats itself every fifth chord (e.g. I IIIb IV# VI I).

4.5 General mapping of intervals in the 12-fold space

Movement through arbitrary intervals in the 12-note space can be plotted on the table given in Fig 12. This table is based on Longuet Higgins' table for intervals in the original space (1962b). The table is an overlay that can be slid around in the 12-fold space, so that the square marked "unison" can be made to coincide with any desired

starting note. The table of 12-fold intervals (Fig 12) repeats itself in all directions in the same way as the array of notes on which it is based (fig 2).

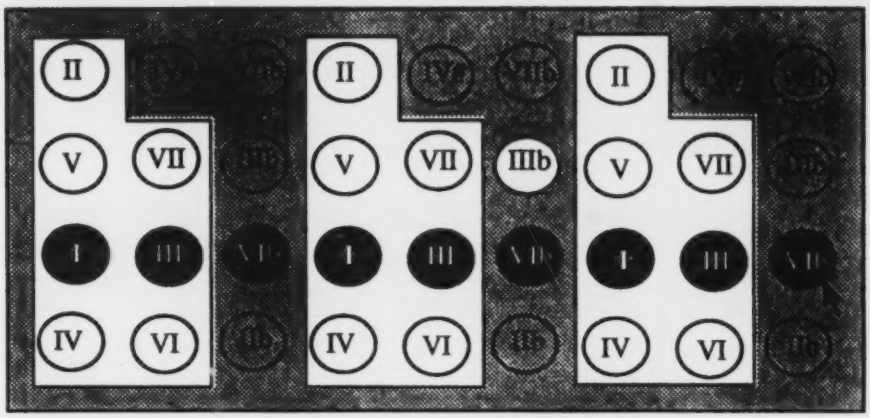


Fig 10 Movement in major thirds unconstrained by key

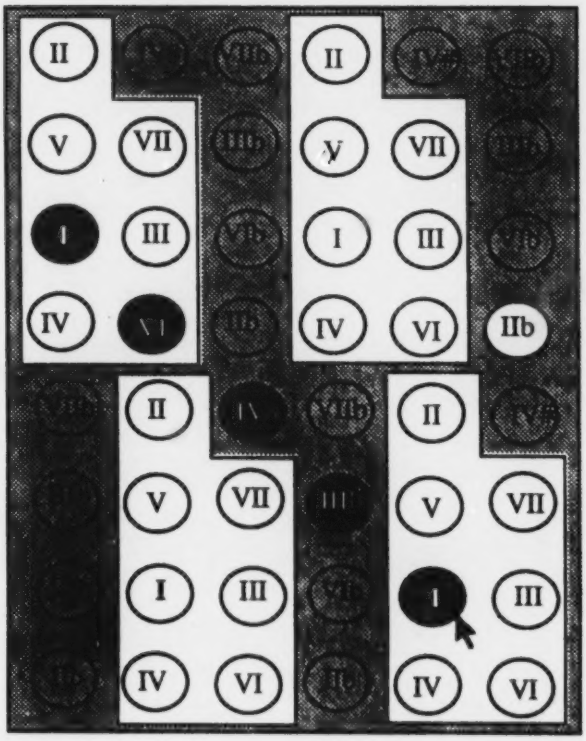


Fig 11 Movement in minor thirds unconstrained by key

Our interest is the relative convenience of the different gestures involved in playing the different intervals in Harmony Space. Ideally, we would like all intervals between the 11 notes of the chromatic scale (under octave equivalence) to be playable by a *single step* in the appropriate compass direction without a discontinuous jump. This would allow novices to play *arbitrary* chord sequences by a simple change of direction of movement, without the need for discontinuous leaps. In the 12-fold space, there are only twelve distinct intervals (under octave equivalence), as against thirty five intervals in Longuet-Higgins' (1962b) table of natural intervals in musical use. The intervals form complementary pairs (two intervals are complementary if their successive application to a pitch yields unison or an exact number of octaves - e.g. major sixths and minor thirds). Complements can be found in the table diametrically opposite each other - i.e. by moving in the opposite sense along a radial axis. Nine of the possible 12 intervals in the 12-fold space can be specified by single vertical, horizontal or diagonal steps to *adjacent* cells in the table. Only the following three intervals involve movement to *non-adjacent* cells: the minor seventh and whole-tone (complementary to each other) and the tritone (complementary to itself). Due to the repeating pattern, there are three choices, more or less equidistant (depending on the metric), for representing each of these three intervals gesturally (fig 12). Fortunately, tones and tritones occurring naturally in a given key (i.e diatonic tones and tritones), can be made, in effect, a single continuous gesture away simply by making notes not in the key window 'out of bounds'. All that is needed is for trajectories to "skip" automatically over chromatic areas between diatonic key-window 'islands'. [11]

Given this arrangement, only progressions in chromatic intervals of tones and tritones (and their complements) ever require a discontinuous gesture (as opposed to a move to an 'adjacent' note) in order to be played with Harmony Space. Of course, discontinuous ways of playing intervals can be used anyway. Each different gestural representation might be viewed as emphasising different ways of looking at the same interval (e.g. a tone as two semitones steps, or as two steps up of a fifth, etcetera). Harmony space allows the different ways of playing the intervals to be explored and discovered - and the different ways of viewing these intervals can be used to make compositional capital.[12] Note that the alternative ways of playing the intervals are

[11] See appendix 2 for a way of implementing this idea.

[12] For example, Bach and Vivaldi sometimes seem to explore the use of different size interval

formally indistinguishable, since intervals of the same name really are the same intervals in this space.

Up a tritone 6 *(Down a tritone)*	*Up a minor seventh* 11 *(Down a tone)*	*Down a minor seventh* 2 *(Up a tone)*	*Up a tritone* 6 *(Down a tritone)*	*Up a minor seventh* 11 *(Down a tone)*
	Up a minor third 3 (Down a major sixth)	Up a perfect fifth 7 (Down a perfect fourth)	Up a major seventh 11 (Down a semitone)	
	Up a minor sixth 3 (Down a major third)	Unison 0 (Octave)	Up a major third 4 (Down a minor sixth)	
	Up a semitone 1 (down a major seventh)	Up a perfect fourth 5 (Down a perfect fifth)	Up a major sixth 9 (Down a minor third)	
Down a minor seventh 2 *(Up a tone)*	*Up a tritone* 6 *(Down a tritone)*	*Up a minor seventh* 11 *(Down a tone)*	*Down a minor seventh* 2 *(Up a tone)*	*Up a tritone* 6 *(Down a tritone)*

Fig 12 Table of intervals in 12-fold space

4.6 Summary of the "dynamics" of harmony

We have now given an illustrated overview of how chord and bass (or melody) progressions can be described and controlled in Harmony Space. Root progressions have been represented as *harmonic trajectories*. Different *spatial directions* and observance or non-observance of *key constraints* have been used to distinguish different progressions. *Tonal and modal centres* have been represented as *goals* for trajectories. Direct spatial metaphors have been provided for common musical intuitions such as that a tonic is about to be sounded, that it will be reached in just so

steps played in regular progressions by different voices to reach strategic targets at carefully timed intervals. This notion is illustrated (in quite other terms) by (Rutter, 1977, page 17).

many steps and so forth. Notice that the ability to represent the staple diatonic progressions as goal-oriented straight lines is a gain resulting from our decision to use the 12-fold space, and simply is not possible in the original space. Straight line trajectories in the original space represent wild modulations.

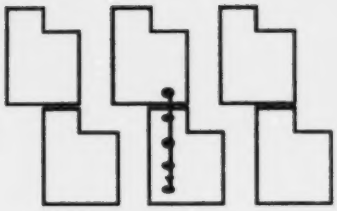
4.7 Harmonic Plans

While we have described some of the building blocks of harmonic progression, composing an interesting chord progression involves far more than stringing together such blocks. One way of describing progressions at a higher level involves musical planning which is one of the matters investigated in Part III.

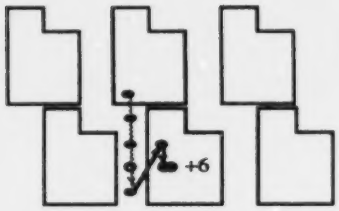
5 Analysing music informally in Harmony Space

So far we have shown that Harmony Space can provide economical means of representing and controlling various low level aspects of harmony, but it turns out that Harmony Space can also be used to illuminate some higher level aspects of harmony in particular pieces. Longuet-Higgins (1962) has analysed Schubert, and Steedman (1972) has analysed Bach in the original space. In both cases, this involved assigning each note a tonal function, much as in conventional harmonic analysis, but we analyse in a different spirit here. To illustrate this, fig 13 gives the chord sequence of the jazz standard "All the things you are" (Kern-Hammerstein). For clarity, only the root of each chord is indicated, and chord alteration is indicated by annotation where the quality departs from the default quality for a seventh chord in the prevailing key. Note also that in fig 13, immediate repetitions of chords with the same root are shown as slightly staggered, overlapping dots.

Let us examine the chord sequence in terms of Harmony Space concepts and try to find a rationale for the chord sequence. Viewed from this perspective, the song appears to break into a small number of recognisable harmonic manoeuvres. In the first eight bars, the sequence begins on a VI chord and makes a trajectory towards the major tonal centre (jazz sequences in minor keys are very rare (Mehegan, 1959)). But the sequence plunges on past I at the fourth bar onto IV at the fifth bar. The song at this point is in danger of breaking a convention for the dialect.



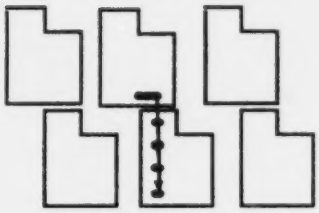
bars 1-5
Starting key I



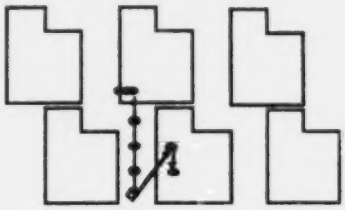
III

bars 6-8
Modulates up a major third in bar 6
(Arrow shows direction of key window shift)

V

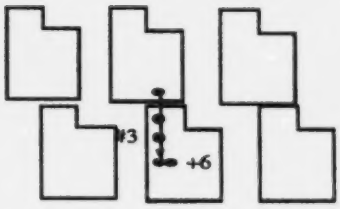


bars 9-13
Modulates up a minor third in bar 9

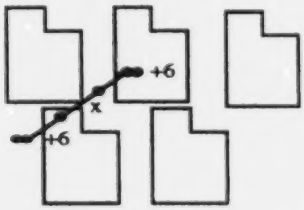


VII

bars 14-15
Modulates up major third in bar 14

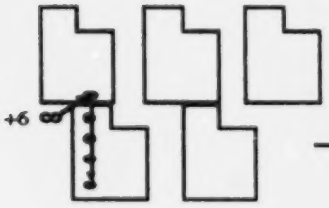


bars 16-20



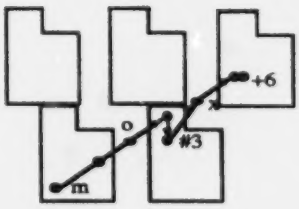
bV

bars 21 - 24
Modulates down a minor third in bar 21



I

bars 24-28
Modulates up a major third in bar 24



bars 29-36

(Chord sequence reproduced from Mehegan (1959)). 'All the things you are' by Kern-Hammerstein © Chappell & Co.,Inc and T.B. Harms Co.

(Ab) VI / II / V / I / IV / (C) V / I / I+6 /
(Eb) VI / II / V / I / IV / (G) V / I / VI / II /
(G) V#3 / I / I+6 / (E) II / bIIx / I / I+6 /
(Ab) VI / II / V / I / IV / IVm / III / bIIIo / II /
(Ab) V#3 bIIx / I+6 / I+6 //

Fig 13 "All the things you are" traced in the 12-fold space

In the chord sequences of jazz "standards" there is a convention that we normally expect to reach a tonal goal when we hit a major metrical boundary at the end of each line (normally eight bars). Subjectively, it is as though the chord sequence has been aimed at a tonal goal but has overshot it in some way. The resolution of this apparent violation of a convention becomes in effect the harmonic motif of the whole piece - we move the 'goalpost'. This is achieved by a timely transient modulation allowing the progression to reach the tonal goal at the metrical boundary. Note that it is common in this dialect where all chords are routinely played as sevenths to emphasise arrival at the tonic (with the subjectively 'restless' or unstable chord quality major seventh) by repeating it as a more stable major sixth. Note also that in fig 13, to make it clearer how chord roots before and after a modulation are related, hollowed out dots are used to recapitulate chords played *before* a given key window move.

In Harmony Space, the 'moving goalpost' metaphor is demonstrated literally. The environment physically shows (fig 13) the 'goalpost' in the form of the key window being moved sideways so that the tonal cycle of fifths arrives at the goal or tonal centre just before the metric boundary is reached. A manipulation of the listener's expectations based on this idea seems to lie at the heart of the progression as a whole. After the move has been performed once, a transient modulation turns the tonic at the end of the first eight bars into a VI. This sets the stage for the moving goal post 'game' to be played again in the second line of the piece. The first time, the 'moving goalpost' may have been quite unexpected, but the second time we are more likely to expect it and if so, we have our expectations confirmed. The third line is split into two parts. In the relevant jazz dialect the progressions involved are more or less stock material, labelled by jazz musicians as "turn-arounds". In Harmony Space these are seen as short harmonic trajectories along the fifths and chromatic axes respectively to temporary tonics. (There is a transient modulation before the chromatic trajectory, which we will not try to explain here, but which seems to function partly to position the final chord of the third line ready for the first chord of the last line.)

As the final line of the sequence begins, it appears to be a repetition of the first line,

so the listener may well expect the 'moving goalpost' manoeuvre to happen again. This time, the expected transient modulation does not happen, and the chord sequence continues on. (In fact, a chromatic progression is used to substitute for part of the cycle of fifths progression.) Hence the last line requires 13 bars, rather than the normal eight bars for the chord sequence to reach the tonal centre and conclude the sequence. It is as though the song originally fooled us with an unexpected manoeuvre, performed the manoeuvre again to make sure we grasped it, and then finally fooled us again by *not* performing the manoeuvre where expected it.

The broad outlines of this informal analysis, if not all the details, should be communicable to a novice with little musical knowledge but experience of Harmony Space. It is important to note that a visual formalism is not being proposed as a substitute for listening. It is being suggested that an animated implementation of the formalism linked to a sounding instrument controlled by simple gestures may allow novices without instrumental skills and without knowledge of standard theory and terminology to gain experience of controlling and analysing such sequences. But such an environment might well be a good place to learn music theory if the novice desired.

6 Conclusions

We have presented a theoretical basis and a detailed design for a computer-based tool, "Harmony Space", designed to enable novices to compose chord sequences and learn about tonal harmony. We have investigated the expressive properties of the interface in detail. Detailed conclusions for this chapter together with discussions of the educational applications of Harmony Space, notes on related interfaces and comments on the limitations of the design, can all be found at the end of Chapter 7 (sections 6-10).

Chapter 6: Adapting Longuet-Higgins' theory for Harmony Space

"The theory described here [Longuet-Higgins (1962) theory] is so simple that it is hard to believe that it has never been formulated before. The facts that it accounts for have been known at least since the time of Morely (1597).... Some early formulations come rather close. Weber (1817).. produced a similar diagram. In his case, however, the diagram was of keys rather than their notes.

Steedman (1972) page 112

"...the gulf between the logical science of acoustics and the empirical art of harmony has never been convincingly bridged..The theory of the generation of chords and harmony from the harmonic series .. must be discarded once and for all." *Groves dictionary , entry on 'Harmony' Vol II (Colles, 1940).*

".. it shocks me a little that music theorists should be, apparently, so ignorant of the two dimensional nature of harmonic relationships.. the whole thing follows relentlessly from first principles" *Longuet-Higgins (1962) page 248.*

In this chapter, we outline the educational and human-computer interface design considerations that led to the use of an adapted version of Longuet-Higgins' theory in the design of Harmony Space. The structure of the chapter is as follows;

1 Longuet-Higgins' non-repeating space
- 1.1 Key areas, modulation and chords
- 1.2 Minor and diminished triads
- 1.3 Chord sequences

2 Comparison of the two versions of Longuet-Higgins' space

3 Direct manipulation and design choices

4 Use of the 12-fold representation in Harmony Space

5 Conclusions

We begin by looking at the structure of Longuet-Higgins' original, non-repeating representation in more detail.

1 Longuet-Higgins' non-repeating space

The non-repeating space is by definition an array of notes arranged in three dimensions, with ascending octaves, perfect fifths and major thirds on the three respective axes (fig 1)[1]. For reasons explained earlier, the octave dimension is usually disregarded.

1.1 Key areas, modulation and chords in the non-repeating space

As the axes are traversed, double sharps and double flats are encountered, in contrast with the twelve-fold space. Despite the repetition of note names, notes with the same name in different positions are *not* identical, but notes with the same name in different possible key relationships.

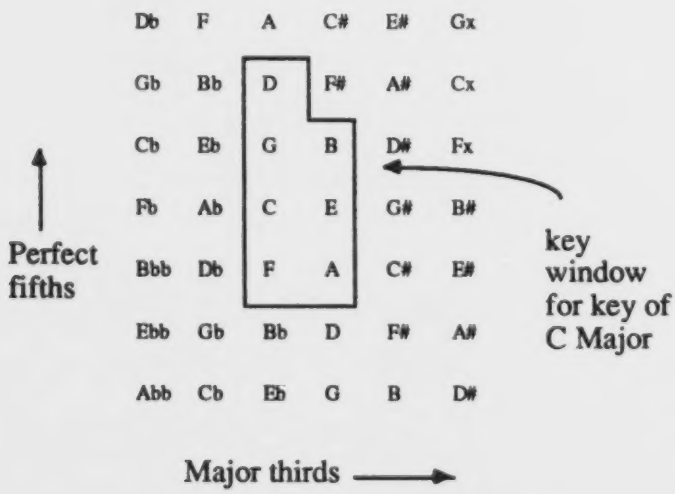


Fig 1 Longuet-Higgins original non-repeating space
(diagram adapted from Longuet-Higgins (1962))

[1] Fig 1 repeated from chapter 5 for convenience.

An illustration may help to clarify the distinction. In a melody in the key of C major, for example, the note A must be represented in different positions depending on whether it is part of a harmony (actual or implied) using an A minor triad (A C E - fig 2), or whether it is part of a harmony using the D major chord (D F# A - fig 3). In the first case, the A is identified with the note of the same name found in the key window in its C major position (fig 2). In the second case, the A is identified with the A that would be in the key window if it were in its G major position (fig. 3). This demonstrates that the spatial position of each note in Longuet-Higgins' original representation encodes not just its pitch class but also the possible keys to which it belongs[2] (whether by a definite modulation or simply by a momentary harmonic "borrowing").

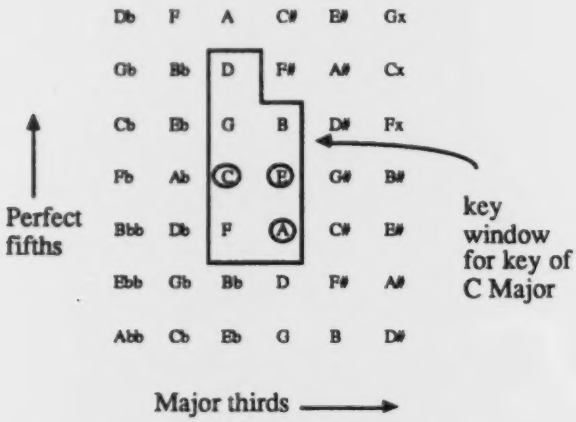


Fig 2: Triad of A minor in the key of C major in the non-repeating space

In the non-repeating space, instead of there being only twelve distinct major tonics, each possible position of the key window (from a potentially infinite number of positions) theoretically represents a distinct key.

---

[2] Interestingly, much the same information can be encoded for each note in the apparently less expressive 12-fold space. Similar information can be encoded using the 12-fold representation by associating the following information with each note; an ordered pair consisting of a spatial position for the note and a spatial position for the key window. Under this scheme, the position of the key window for a given note may be different from the position of the key window for other simultaneously occurring events - for example, when one of notes is "borrowed' from a foreign key. This scheme is less ambiguous about key relationships than the non-repeating space, since a location in the non-repeating space encodes for several *possible* keys rather than just one key.

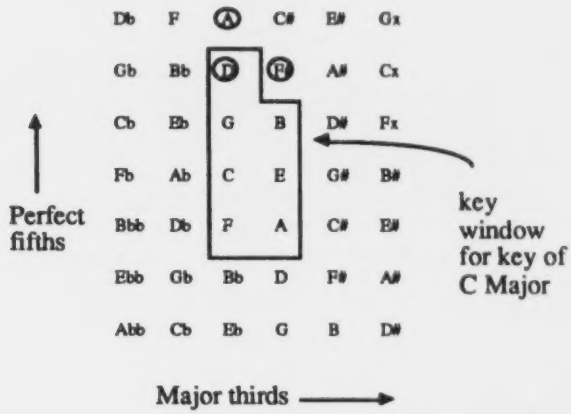


Fig 3: D minor triad in the key of C major in the non-repeating space

Major triads are represented in the non-repeating space in a way that is similar in many, but not all, respects to the way they are represented in the 12-fold space. As before, major triads correspond to triangular shapes (fig 4). Three distinct notes are maximally close in the space when they form a major triad. Major triads can be distinguished from other maximally compact three element shapes in that they are generated by selecting a note and choosing one more note along both axes in the *outward-going* (i.e. intervalically ascending) sense. There is a clear spatial metaphor for the centrality of the major tonic - the tonic triad is literally the central one of the three major triads that generate any major key.

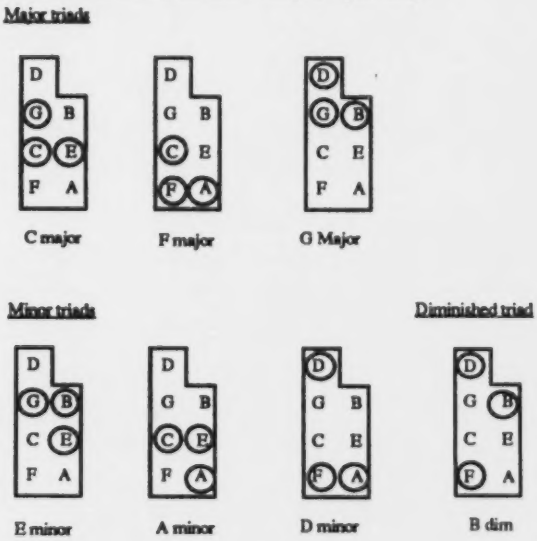


Fig 4

Triads in C major in the non-repeating space

1.2 Minor and diminished triads

Differences between the representations come to light when considering the minor and diminished triads. Two of the three minor triads (III and VI) correspond to rotated triangular shapes (fig 4). These minor triads are, as in the 12-fold space, maximally compact three-element objects. However, the third minor triad (II) is a different shape from that of the other minor triads and its constituent notes do not lie contiguously in the key window (fig 4). The shape for the chord of D minor in the key of C, for example, is a different shape from the chords E minor and A minor in the key of C. This non-uniformity of shape in the minor triads is inescapable in the non-repeating space, since, as we have already seen, spatial position codes for possible key relationships. To use the same shape for the D minor triad would be to misleadingly suggest a transient modulation to the subdominant or a related key. Furthermore, since the minor triads do not form a uniform group of contiguous objects, the spatial metaphor for the perceived centrality of the minor tonic falls down in the non-repeating space. To sum up, the differences in the way minor triads are represented in the 12-fold and non-repeating spaces are as follows. Firstly, in the 12-fold space, but not in the non-repeating space, there is a one-to-one mapping between shape and chord quality for both major and minor triads. In the non-repeating space, the minor triads have more than one shape. Secondly, the 12-fold space, but not the non-repeating space provides a single metaphor for the centrality of both major and minor tonal centres.[3] A final difference between the representation of triads in the two spaces is that all triads in the 12-fold space are connected [4], whereas the scale-tone diminished chord (VII) in the non-repeating space has the non-connected shape shown in fig 4.

1.3 Chord sequences in the non-repeating space

In the 12-fold space, chord progressions in cycles of fifths, scalic and chromatic

[3] This 'imbalance' between the major and minor modes in the non-repeating space is underscored by another relationship (occurring only in the non-repeating space) that emphasises the centrality of the major tonic (but not of the minor tonic). The relationship is as follows. The tonic position is the only position from which all other notes in the key can be reached in a maximum of two moves - horizontal, vertical or a mixture (Sloboda, 1985).

[4] i.e. all notes in the set are horizontally, vertically or diagonally immediately adjacent to at least one other note in the set.

progressions all appear as straight lines. In the non-repeating space, this is not the case, unless wild modulations are taking place. Scalic and chromatic progressions are represented by a sequence of non-contiguous steps in constantly changing directions. Fig 5 compares a standard chromatic progression in the 12-fold space with its equivalent in the non-repeating space. (The precise details of the chromatic progression in the non-repeating space are not important for our purposes.)

2 Comparison of the two spaces

The most important differences, for our purposes, in the way that basic harmonic relationships are represented in the 12-fold space and the non-repeating space can be summarised as follows.

- In the non-repeating space, notes with the same name in different positions have different key implications. In the 12-fold space, there are only 12 distinguishable notes in a repeating pattern.

- In the 12 fold space, there is a one-to-one correspondence between chord quality and shape for the scale-tone triads. There is not in the non-repeating space.

- In the 12-fold space, the shapes of all of the triads are compact. In the non-repeating space, the chords II and VII do not have connected shapes.

- The 12-fold space, there is a simple metaphor to account for which of the major and minor triads are aurally perceived as 'central' - the non-repeating space only does this for the major triads.

- Straight lines and simple patterns in the 12-fold space correspond to the basic chord progression of western tonal harmony; cycles of fifths, scalic progressions, chromatic progressions and cycles of thirds. In general, the same chord progressions in the non-repeating space are complex patterns. Straight lines in the non-repeating space correspond to extreme modulations.

3 Direct Manipulation

We have now assembled nearly all of the information needed to explain why we mapped Longuet-Higgins' onto a 12-fold note set when designing Harmony Space. The final idea to be considered is the notion of 'direct manipulation'.

Shneiderman (1984) noted the fact that some interactive systems seemed to promote "glowing enthusiasm" among users, in contrast with the more widespread reaction of "grudging acceptance or outright hostility". Shneiderman identified what seemed to be a central core of ideas associated with the 'felicitous' systems, and labelled the systems that used these ideas 'direct manipulation systems'. The three main ideas were (using paraphrases of Shneiderman (1984) and (Hutchins, Hollan and Norman, 1986),

- continuous representation of the objects of interest.

- replacement of complex command language syntax by direct manipulation of the object of interest through physically obvious and intuitively natural means, including, where appropriate, labelled button presses.

- rapid, reversible, incremental actions whose impact on the object of interest is visible.

Many benefits have been claimed for direct manipulation, such as (paraphrasing (Hutchins, Hollan and Norman, 1986)

- quicker learning by novices, especially where demonstrations are available,

- knowledgeable intermittent users can retain operational concepts,

- error messages are rarely needed,

- users can see if their actions are furthering their goals, and if not, change what they are doing,

- users have reduced anxiety because the system is comprehensible and actions are reversible.

It can be hard to be precise about what does and does not constitute an example of a direct manipulation system, and there is not universal agreement about which of the claims are borne out and to what extent.

4 Use of the 12-fold representation in Harmony Space

We now explain (from an educational and human-machine interface design perspective) why we mapped Longuet-Higgins' space onto 12 notes when designing Harmony Space. We could have constructed a direct manipulation learning environments for harmony using either version of Longuet-Higgins theory, but as we have just seen, it is only in the 12 -fold space that the most common chord progressions correspond to *physically obvious and intuitively natural ways* of moving objects in the space - i.e. straight lines. In the non-repeating space, a descending chromatic scale, for example, properly corresponds to a unique intricate path around a fixed key window that would not be physically obvious and intuitively natural at all, even to an expert musician. Since the ability for novices to manipulate common chord progressions in a natural way is one of our aims, the 12-fold space confers a relevant advantage. Traded off against this advantage, the 12 fold space ignores the distinction between, for example notes such as G and Abb. However, for novices at the level and for the purposes that we are considering, this distinction is one that we would rather not introduce anyway. There is a second, related reason for preferring the 12-fold to the non-repeating space as the basis for a direct manipulation environment. In an environment based on the non-repeating space, the notational distinctions (e.g. between Gb and F#) would be available, but it would be easy for a novice to make the *wrong* notational distinctions continually, for example when playing a given melody. A great deal of musical knowledge can be needed to choose the right version of the note, knowledge that novices could not be expected to possess.

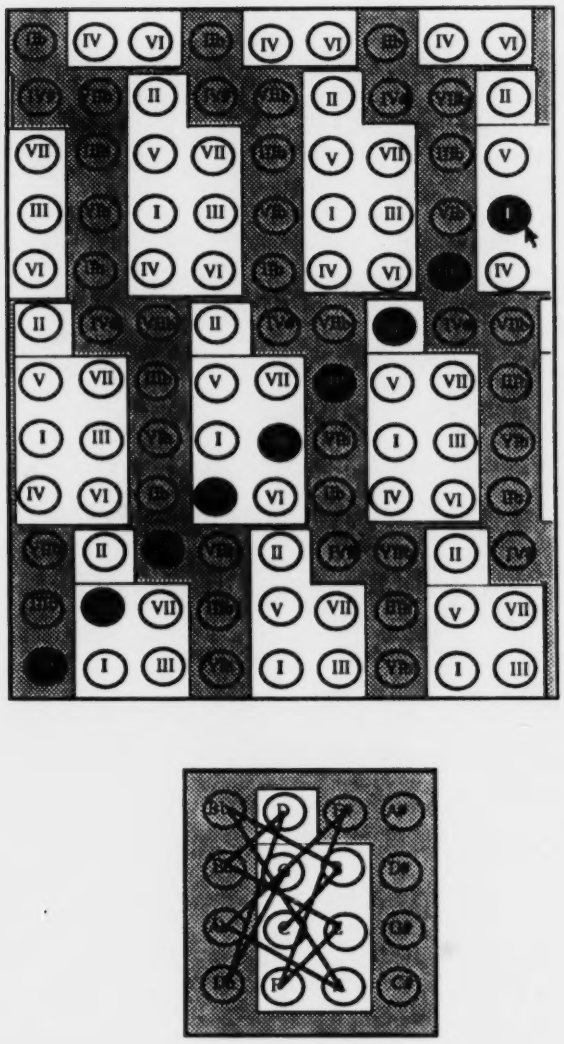


Fig 5

Below: A chromatic progression in non-repeating space

Above: Part of the same chromatic progression in 12-fold space

Neither could the interface embody the knowledge, since we do not yet know how to formalise it in the the general case (Steedman, 1972). The notationally 'wrong' notes would sound indistinguishable from the notationally 'right' notes. Hence, an interface based on the non-repeating space would introduce a distinction to novices

which they would be likely to continually confuse, yet which the interface would have no power to enforce or teach. The 12-fold space avoids this problem by presenting harmony at a level of detail at which the distinction need never be made.

There is a third set of reasons why it is educationally useful to map the non-repeating space onto the 12 fold space. The designers of the Xerox Star interface (Smith, Irby, Kimball and Verplank, 1982) emphasised the value in interface design of the uniform and consistent representation of the more commonly used entities. This can reduce learning time, lessen the need for memorisation and reduce cognitive load. In line with this approach, we wish to provide uniform and consistent visual analogues for the most commonly used musical entities appropriate to the level of user and task. By the simple measure of mapping the non-repeating space into the 12-fold space, we produce the following results;

- The most commonly used entities (the various chord qualities) exhibit a one-to-one mapping between their most important attribute (chord quality) and their graphical shape. This is a good example of a uniform and consistent representation (Smith, Irby, Kimball and Verplank, 1982).

- The one-to-one mapping between chord quality and chord shape, and the simple contiguous nature of the shapes are both likely to reduce cognitive load when trying to process quickly what is happening on the screen.

- The contiguous, regular shapes of the triads are likely to be easier to recognise than the complex non-contiguous shapes of the non-repeating space.

5 Conclusions

We have now concluded the description of our reasons for mapping Longuet-Higgins' space onto 12 notes for the purposes of designing Harmony Space. The different representations are best suited for dealing with harmony at different levels of description. We have chosen the representation to suit our purposes and prospective clients. The 12-fold representation allows us to describe harmonic relationships completely accurately at the level we have chosen to describe them.

Chapter 7: Harmony Space using Balzano's theory

In this Chapter we show how Balzano's (1980) theory of harmony can be used as the basis for a version of Harmony Space. A prototype of this version of Harmony Space has been implemented. The chapter begins with a presentation of Balzano's theory. We then review the design decisions for Harmony Space in the light of this theory. Next, we discuss the relationship between Longuet-Higgins-based and Balzano-based versions of Harmony Space. The implementation is discussed. In the final sections (which refer to both of the Harmony Space designs together) the following topics are considered: Harmony Space's educational uses; the intended users; related interfaces; the limitations of the implementation; and final conclusions. The structure of the chapter can be laid out as follows;

1 Balzano's theory of Harmony
- 1.1 Basis of theory
- 1.2 Co-ordinate systems for the diatonic scale
  - 1.2.1 The semitone space
  - 1.2.2 The fifths space
  - 1.2.3 The thirds space
- 1.3 Some issues addressed by the theory

2 Review of design decisions
- 2.1 Universal commands and consistency
- 2.2 Core design decisions
- 2.3 Other design decisions
- 2.3 The 'shear' operation

3 Harmony Space using Balzano's theory: design decisions
- 3.1 Core design decisions
- 3.2 Design decisions concerning the display
- 3.3 Universal commands and consistency

4 Expressivity of a version of Harmony Space based on thirds space
- 4.1 Statics of harmony
  - 4.1.1 Fifths axis and semitone axis
  - 4.1.2 Scales and triads
  - 4.1.3 Tonal centres
  - 4.1.4 Visual rule for chord construction

4.2 Representing the dynamics of harmony
5 Analysing a chord progression in the thirds space
6 Relationship between the alternative Harmony Space designs
7 Implementation of Harmony Space
8 Some educational uses of Harmony Space
9 Intended users of the tool
10 Related interfaces
11 Limitations
12 Conclusions
12.1 Design decisions and expressivity
12.2 Uses of Harmony Space
12.3 Longuet-Higgins' theory and education

1 Balzano's theory of harmony

Balzano's (1980) theory of harmony has a completely different starting point from Longuet-Higgins' theory, (mathematical group theory instead of overtone structure) but leads to a remarkably similar representation. Balzano's theory can be seen as addressing the following (among other) problems.

- Why are the notes C and A in the set of notes A B C D E F G (for example) perceived to be central or tonic notes?

- Why are major and minor triads perceived to be fundamental to tonal harmony (as opposed to chords with some other number of notes or interval structure)?

- Why are just, Pythagorean and equal temperament perceived to work in "substantially the same way" (Balzano, 1980)?

1.1 Basis of Balzano's theory

Balzano (1980) points out that *any* pitch set can be considered a cyclic group, by virtue of two observations. Firstly, frequency or (more practically) log frequency can be used to impose a total ordering on the set. (Later we will make use of the successor or "next" relation implicit in this total ordering.) Secondly, octave

equivalence can be used to provide an equivalence relation on the pitch set. The equivalence relation allows all pitches to be given equivalents in each octave; it also allows us to take a set of representatives in the compass of one octave to stand for the entire pitch set. Once the pitch set has been extended by propagating it through the octaves in this way, it is clearly isomorphic to the set of positive and negative integers, and thus isomorphic to C_∞, the cyclic group of infinite order. The set of n representatives from any given octave can be shown to be to be isomorphic to C_n, the cyclic group of order n. If the pitches in one octave are labelled 0 to n-1 (where the note n is an octave higher than the note 0), we say that the pitch set exhibits 'n-fold octave division'. Two complementary and isomorphic ways of mapping the elements and operations for these groups onto the pitch set are possible. According to the first view, the elements are pitches, and the operation of adding two pitches is taken to be equivalent to adding their displacements from some reference pitch as measured in terms of successive application of the successor function mentioned above. According to the second view, the group elements are the set of *intervals* between pitches (sized in numbers of steps), and the operation is taken to be *interval addition*. We will use both views, depending which is appropriate at a given time.

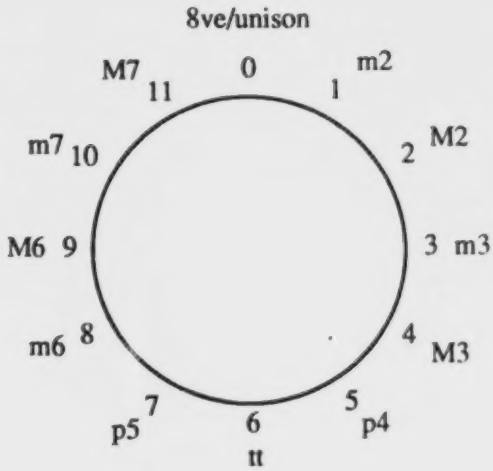
Having established a relationship between pitch sets and cyclic groups, and bearing in mind that western musicians have long agreed to divide the octave into 12 notes, it turns out to be possible to relate tonal harmony to the structure of C_{12} (the cyclic group of order 12). Just before we investigate this, notice the generality of application of the theory. For example, the pitch set involved could be equal tempered, just tempered or irregularly tuned. For practical purposes, we will adopt equal temperament, since this has the advantage that we can specify interval size fully by counting steps between group elements. This will allow symmetry arguments to be used - the set of pitch places will be unaffected by transposition.

1.2 Investigation of co-ordinate systems for the diatonic scale

The key behind Balzano's theory is to investigate all possible well-defined co-ordinate systems for the western pitch set.[1] ('Well-defined' here simply means

[1] This is also the key to Longuet-Higgins' theory, but in terms of an entirely different characterisation of pitch.

that each element in the pitch set must be specifiable and have unique co-ordinates.) Taking a pitch set with a 12-fold division of the octave, let us now informally investigate all possible co-ordinate systems that provide unique co-ordinates for all intervals in terms of successive applications of some interval. The first one-dimensional co-ordinate system is very straightforward: it is simply a semitone grid. (The 'grid' is circular because of octave equivalence.) In other words, an interval of any size less than or equal to an octave can be described trivially as being equivalent to some number (from 0 to 11) of successive steps of 1 semitone. This one-dimensional co-ordinate system gives unique co-ordinates for each interval (Fig 1).

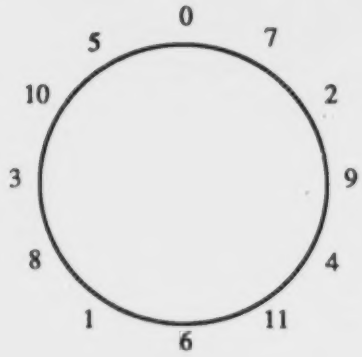


Step size 1 or 11
No of elements in cycle 12
Elements in cycle {0,1,2,3,4,5,6,7,8,9,10,11}
C_{12}

Fig 1 The semitone space. (After Balzano (1980))

An equivalent co-ordinate system is the one we obtain by describing intervals in terms of successive steps of size 11 (remaining in one "octave" since octave equivalence is in force). Steps of size 11 are, in effect, just steps of size 1 in the reverse direction (fig 1). Music theory and group theory are in unusual terminological agreement in calling such pairs of intervals "inverses". The co-ordinate systems that pairs of inverses give rise to are essentially identical, so having identified the existence of such pairs, we shall sometimes use one representative to refer to both in future.

We will now quickly investigate exhaustively the one-dimensional co-ordinate systems of other step sizes for C_{12}. Steps of size 2 (tones), 3 (minor thirds), 4 (major thirds) and 6 (tritones) and their inverses fail even to span the space of all intervals (figs 3 and 4). Successive steps of size 2 or 4 cannot describe intervals of odd numbers of semitones at all (figs 3 and 4). Successive steps of size 3 cannot describe *all* intervals of even numbers of semitones (fig 4). Successive steps of size 6 can only describe intervals of size 6 and 0 (fig 3). The only interval apart from semitones that spans the space of all other intervals is the perfect fifth (step size 7) (and its inverse the perfect fourth - step size 5). This co-ordinate system turns out to provide unique co-ordinates for each interval, since after 11 applications, all intervals have been described exactly once (fig 2). We have now informally demonstrated that in the case of one-dimensional co-ordinate systems (under octave equivalence), only steps of size 1 and 7 (ignoring inverses) have the property of spanning the space of all intervals.

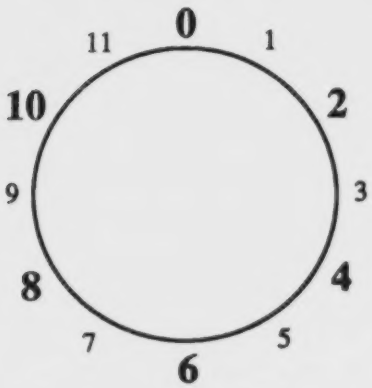


Step size 7 or 5
No of elements in cycle 12
Elements in cycle {0,1,2,3,4,5,6,7,8,9,10,11}
C_{12}

Fig 2 The fifths space. (After Balzano (1980))

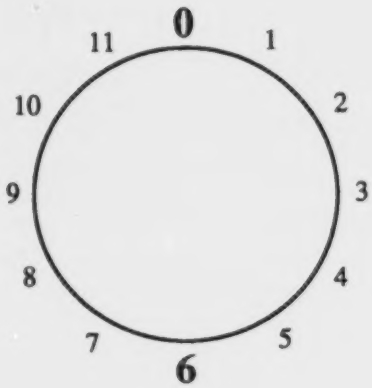
Let us now look to see if any two-dimensional, or higher-dimensional arrays constitute adequate co-ordinate systems for C_{12}. We have already seen that steps of size 2,3,4 and 6 are inadequate in isolation as bases on which to describe all intervals, but perhaps two or more of them could act in combination as adequate co-ordinate systems. Step sizes 2, 3, and 6 are not suitable to use in any combination,

since this would give rise to non-unique sets of co-ordinates for describing a step size 6 (e.g. 2 steps of size 3, 3 steps of size 2). This would violate our requirement for unique co-ordinates for intervals, yielding an unsatisfactory metric.



Steps of tones or major sevenths

Step size 2 or 10
No of elements in cycle 6
Elements in cycle {0,2,4,6,8,10}
C_6



Steps of tritones

Step size 6 only
No of elements in cycle 2
Elements in cycle {0,6}
C_2

Fig 3 Steps of tones and tritones

Without using steps of size 1 and 7 (that we already know can do the job by themselves), of all combinations of two step sizes this leaves only steps of size 4 in combination with steps of size 2, 3, or 6 (ignoring inverses). Combining steps of size 4 with steps of size 2 would lead to a non-unique specification of steps size 4. Steps of size 4 combined with steps of size 6 would be inadequate to specify steps of size three at all. This leaves a sole candidate for a two dimensional co-ordinate system - steps of size 3 in combination with steps of size 4.

A two dimensional grid with 4 steps of size 3 against 3 steps of size 4 spans the space uniquely. The search can be extended to the n-dimensional case and an existence result formally proved in group theory (Balzano 1980), giving us Balzano's central formal result:-

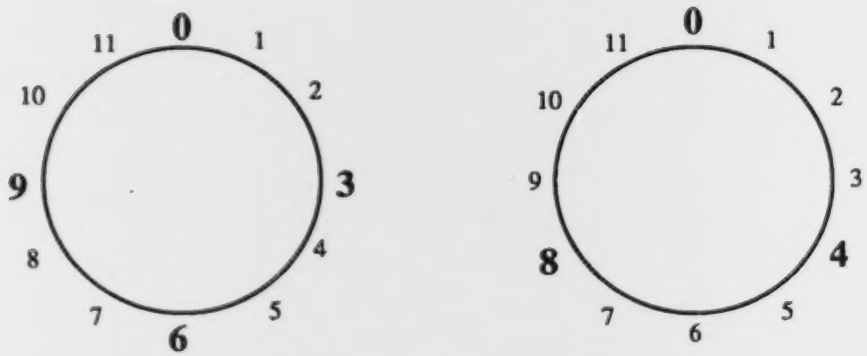
> In any pitch set with a 12 fold division of the octave and octave equivalence, the only co-ordinate systems capable of specifying every interval uniquely under octave equivalence are the following;

0 - 11 minor seconds (step size 1)

0 - 11 perfect fifths (step size 7),

0 - 2 major thirds (steps size 4) and 0 - 3 minor thirds (steps size 3).

No other distinct well-defined co-ordinate systems for C_{12} can be constructed even using arbitrary n-tuples of intervals (Balzano, 1980). The next key move in Balzano's theory is to investigate each of the three co-ordinate spaces in turn, as we will now do.



Steps of minor thirds

Step size 3 or 9
No of elements in cycle 4
Elements in cycle {0,3,6,9,}
C_4

Steps of major thirds

Step size 4 or 8
No of elements in cycle 3
Elements in cycle {0,4,8}
C_3

Fig 4 Steps of major thirds and minor thirds

1.2.1 The semitone space

The semitone space can be represented as in figure 1. Intuitively speaking, 'closeness' in this space corresponds simply to the number of semitones separating two points. Balzano suggests that this measure is closely related to the notion of closeness in melodic motion. We have already seen that both the minor second (step size 1) and the major seventh (step size 11) generate the entire group C_{12} {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11}. It turns out that other elements of the group can be used to generate subgroups. The elements major second (step size 2) and, by symmetry, minor seventh (step size 10) can be used to generate the subgroup C_6,

126

corresponding to the whole tone scale {0,2,4,6,8,10}. The major third (step size 4) generates C_3, corresponding to the augmented triad {0,4,8}, and the minor third (step size 3) generates C_4, corresponding to the diminished seventh chord {0,3,6,9}. The tritone (step size 6) generates C_2, {0,6} and is the only element to be its own inverse element. The group theoretic perspective in the *semitone space* thus provides some new perspectives on a number of traditionally well known harmonic structures, but we must look to the other co-ordinate spaces for the real impact of the theory .

1.2.2 The fifths space

The fifths space (fig 2) proves to be a rich source of insights, of which we will pick out three. Firstly, the cycle of fifths is shown to have a unique status among intervals *irrespective* of the pitch ratio used to represent it. As we have already seen, the perfect fifth (step size 7) and its inverse, the perfect fourth (step size 5) are the only intervals aside from the semitone (and its inverse) that span the entire space of intervals. (This fact is exploited by Pythagorean temperament.) Secondly, the diatonic scale emerges as a natural structure in the fifths space, since it is connected in this space (Fig 5). Thirdly, properties emerge from the representation that, in combination, make the diatonic scale *uniquely* privileged among possible scales as a system for communicating location and movement. (More information about these properties can be found in Balzano (1980).)

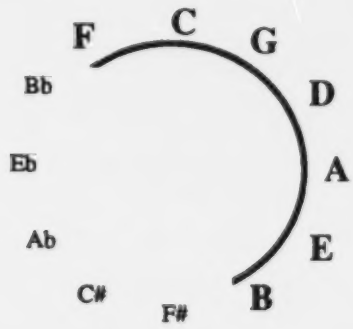


Fig 5 Diatonic scale embedded in fifths space (after Balzano (1980))

1.2.3 The thirds space representation

This leaves us with the final co-ordinate system to explore, the cartesian product C_3

x C_4. The usual graphic representation of this space is given in Fig 6. Notes are indicated by semitone numbers, and the diatonic scale is picked out as a parallelogram.

0	4	8	0	4	8	0	4	8
9	1	5	9	1	5	9	1	5
6	10	2	6	10	2	6	10	2
3	7	11	3	7	11	3	7	11
0	4	8	0	4	8	0	4	8
9	1	5	9	1	5	9	1	5
6	10	2	6	10	2	6	10	2
3	7	11	3	7	11	3	7	11

Fig 6 Balzano's thirds space (After Balzano (1980))

It should be evident that superficially, this can be viewed as the Longuet-Higgins space having undergone a shear operation. This may be more obvious looking at essentially the same diagram displayed slightly differently (fig 7). Leaving aside for a moment the exploration of the thirds space in detail, we can already indicate how the Balzano space addresses the issues noted at the beginning of this section.

• *Centrality of major and minor tonics.*

In effect, we get the same characterisation as in the case of the 12-fold version of the Longuet-Higgins' space - the major and minor tonic triads are revealed as spatially central in the major and minor triads respectively (fig 8).

• *Why are major and minor triads perceived to be fundamental to tonal harmony?*

Balzano (1982) points out that prior to the growth of the triadic basis of music, six of the possible tonics in the seven note scale had been extensively

used as tonics, the tonics being marked not structurally by the scale, but by repetition, emphasis etcetera. In the fifths space, no tonic is particularly marked out by its location. Balzano speculates that the emergence of two structural tonal centres and the adoption of triadic textures were mutually dependent. Balzano (1980) quotes Wilding-White (1961) to the effect that the central notion behind a tonic is not a central note, but a central triad. Hence, the importance of triads may be due to the way in which their use gives rise to structural tonal centres

• *Why are Just, Pythagorean and equal temperament perceived to work in "substantially the same way"*

All three tunings on a fixed tuning instrument have exactly the same configurational (i.e. group theoretic) properties. Balzano (1980) says,

> When we look at the Just, Pythagorean, and Equal Temperament tunings of our familiar major scales, it is evident that there is an important sense in which they all work in substantially the same way: we don't have to learn to compose or hear separately for each one. The ratios, which are different in all three tuning schemes, do not really affect this fundamental commonality. (Balzano, 1980, page 83)

Balzano's theory also provides answers to other fundamental questions about tonal harmony, such as why the octave is divided into twelve pitches; why a seven note scale is selected from twelve pitches, and what other pitch systems are likely to be psychologically interesting from a harmonic point of view (Balzano, 1980)).

It is sometimes objected that the Balzano's theory applies only to western music, and not to music based on various ethnic scales. There is evidence that Balzano's theory is much more universal than this objection might suggest. Sloboda (1985) gives three reasons why 12-fold scales may be much more widely used in other musical cultures than at first appears. Firstly, pentatonic scales are simple subsets of the diatonic scale. Secondly, the diatonic scale is widespread both geographically and historically, although exact tunings may vary (Kilmer, Crocker and Brown, 1976). Thirdly, Sloboda (1985) cites Jairazbhoy (1971) to the effect that where scales are theoretically not 12-fold, such as the Indian 22-fold *sruti* scale, they may be used in

a 12-fold manner in practice.

3 Harmony Space using Balzano's theory: design decisions

3.1 Core design decisions

We can now bring across all of the design decisions from the core abstract design of the 'thirds/fifths' Harmony Space tool of chapter 5, the only change being, in effect, a graphic shearing of the screen layout. We will now review the major design decisions. We have a direct manipulation environment using the thirds space that sounds in real time, with user-selectable chord arity. The chords have default qualities indexed by the degree of the mode on which they are based. The key window is visible and movable. The default qualities are made comprehensible by use of the key window as a visible constraint mechanism. At least two distinct pointing devices are required, one for moving notes and one for moving the key window. Extra pointing devices can be used, if required, to simultaneously control multiple voices, perhaps in different registers and perhaps with different arities. Different voices can be controlled by different users. Foot switches or pointers can be used to change chord-arity, enforce or relax key constraints, overide chord quality and assign rhythmic figures to voices.

3.2 Design decisions concerning the display

As in the previous version of Harmony Space, bearing in mind that we will be looking at trajectories of chord and bass progressions in the space, it makes sense for us to explicitly multiply the key window like a wall-paper pattern, in the same way that this has been done for the notes. To emphasise similarities with the Longuet-Higgins 12-fold space, we graphically pick out the diatonic scale using a box design (fig 7), rather than the equally possible parallelogram (fig 6).

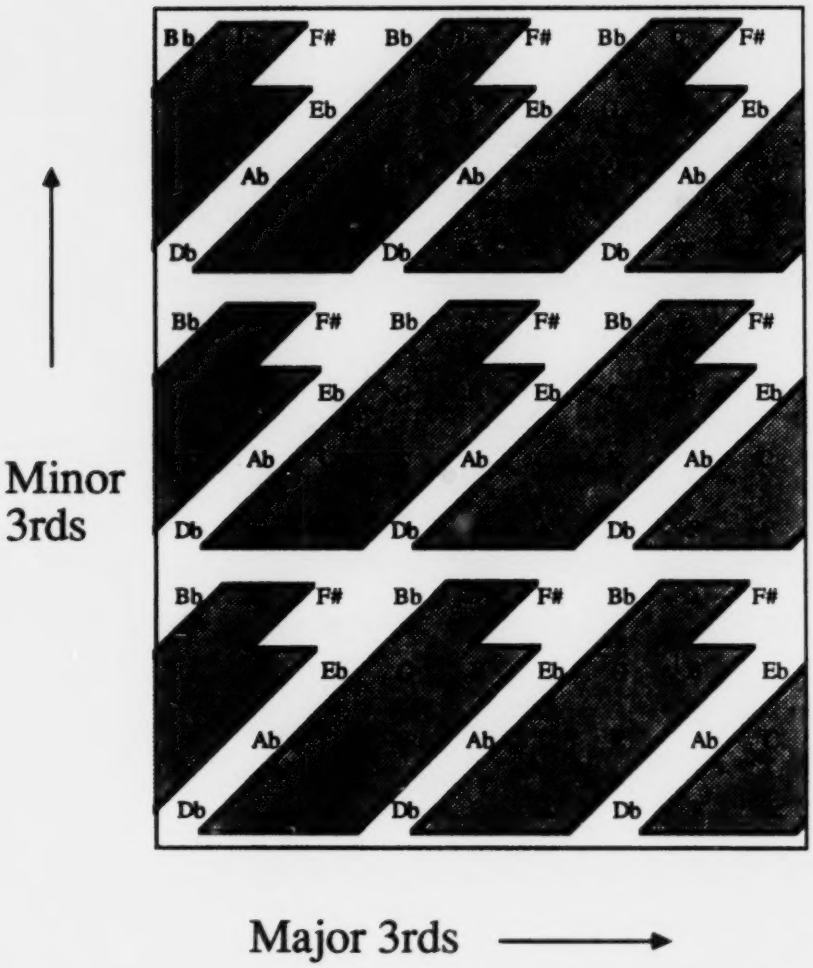


Fig 7 Harmony Space using thirds representation

We slice up the diatonic areas (which could alternatively be left as continuous zig-zagging strips) into boxes (fig 7). This 'slicing' is a conscious design choice. The motivation for it is to make the display easier to "chunk" visually, and to make mental comparisons with the Longuet-Higgins non-repeating space easier. The slice as we have positioned it may help emphasise the stacking of the major thirds and the centrality of the major tonic, but equally, it may slightly obscure the centrality of the minor tonic. (The strips could be sliced up horizontally between II and V instead of between VI and II, which would tend to give the opposite emphasis.) The placing, or absence of the slice in the display has no theoretical significance (see footnote 2 in chapter 5). Note spelling has no significance, following vernacular usage. Sharps

could be used consistently, or neutral semitone numbers could be used in place of note names.

3.3 Universal commands and consistency

In a design for Harmony Space based on Balzano's theory, the semitone and fifths space could have been chosen to feature equally prominently with the thirds space. For example, the current key could be controlled and displayed by the position of a dot in a circular display representing the fifths space (like fig 2). The thirds space direct manipulation area could be held fixed during modulation. Diagrams of all three representations (the semitone space, the fifths space and the thirds space) could be used both as display and direct manipulation control areas for chords or notes. This might make interesting alternative version of Harmony Space. However, we have decided to opt for a small number of universal gestures (commands) that work consistently for as many functions as possible. Thus the same diagonal displacement represents a note, a triad or a key moving through a perfect fifth, depending on which (easily distinguishable) entity the user has hold of. This has advantages of uniformity and consistency as advocated by Baecker and Buxton (1987, page 605). In our case, the uniformity allows the same gestures, localised in the same part of the screen, to be used to control all functions. It also allows all of the relationships to be visualised in one place. The extra cognitive load, memory imposition and discontinuous gestures involved in moving between sub-displays with different styles of command gesture are avoided. This uniformity is made possible without distorting the theory by the fact that the thirds space representation subsumes the semitone and fifths spaces, which appear as diagonals in the thirds space (fig 7). [2] However, this is not to argue that there is any "right" answer to this design issue. A different argument might be that multiple alternative representations are educationally desirable. It is likely that both designs would have advantages for different purposes. We will concentrate on the version with a single uniform display. We have now arrived a Harmony Space interface based on the Balzano theory whose appearance is as in fig 7. We will now run through a compressed exposition of the elements of tonal harmony to see how they appear in the new design (of section

[2] The only aspect of the one-dimensional spaces the two-dimensional space fails to show explicitly is their circular aspects. But we can imagine addressing even this shortcoming with toroidal or cylindrical representations on screen that rotate appropriately as objects on their surface are moved about.

3.1), noting how familiar expressive features are emphasised in different ways. As we will see in a moment, the results are slightly different from what a simple shear operation might lead us to expect. In any case, we have no way of getting an intuitive feel for which elements of tonal harmony may appear clearer in one representation or the other without explicitly inspecting them.

4 Expressivity of a version Harmony Space based on thirds space

4.1 Statics of harmony

The two axes of the display are the major and minor third (fig 7). As we have already noted, because of the circular nature of the chromatic scale under octave equivalence, the two dimensional grid (fig 7) really describes the surface of a torus. However, to make relationships easier for the eye to read, the torus has been unfolded and repeated like a tile pattern. Duplicate note names represent exactly the same note (unlike in Longuet-Higgins' non-repeating space).

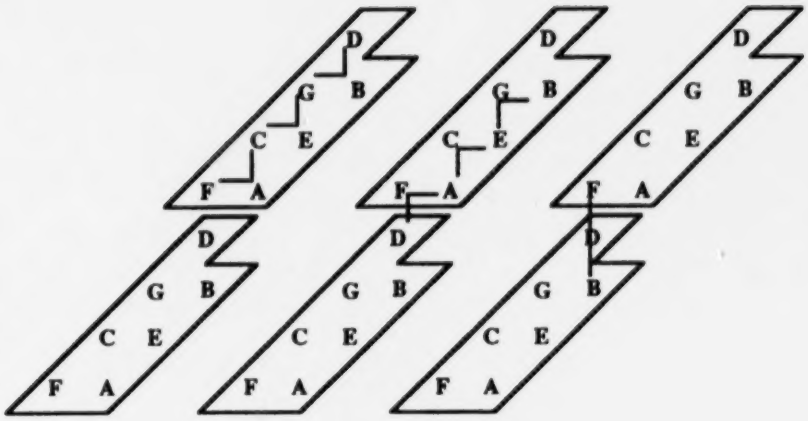
4.1.1 Fifths axis and chromatic axis

From an educational point of view, two important dimensions that western musicians constantly make use of (and novices need to learn about) have simple physical, spatial interpretations in Balzano's space. "North-East" diagonals correspond to cycles of fifths and the "South-East" diagonals are chromatic scales. The thirds space can be viewed as subsuming the two one-dimensional representations.

4.1.2 Scales and triads

The notes of the diatonic scale clump together into a compact window containing no non-diatonic notes. Moving the key window corresponds to modulation. If we slide the window up the diagonal cycle of fifths axis starting from the key of C, we will lose an A and an F at the bottom of the window, but regain the A and capture an F#. Clearly this corresponds to modulating to the dominant. The *only* maximally compact three element shapes that fit into the diatonic window are none other than the major and minor triads. (Fig 8) Either set of three triads span the diatonic space. Every degree of the diatonic scale has (selecting notes in a positive-going direction)

a major or minor triad associated with it , except for VII (the leading note) with its associated diminished triad (fig 8).



Major triads in C major I, IV and V	Minor triads in C Major II, III and VI	Diminished chord VII

Fig 8 Triads in the thirds space

4.1.3 Tonal centres

If we consider the diatonic scale as swept out by translating a major triad by two successive moves along the cycle of fifths diagonal, then we can see that the triad whose root is the tonal centre of the major mode is visually (and computationally) the central of the three major triads (fig 8). The strongly related (and harmonically next most important) primary triads are immediately to either side of the tonic triad. The triad of the minor mode is centrally placed between the closely related secondary minor triads. This gives us a very simple story to tell to novices about the nature and location of the major and minor tonal centres.

4.1.4 Visual rule for chord construction

It turns out that the use of the thirds space may give us a gain, from an educational viewpoint, in the way that the building blocks of harmony (scale tone triads and sevenths, etcetera) can be seen to be built up. In the case of the fifths/thirds version of Harmony Space, the rule for choosing the next note as chords of higher arity were built was not as obvious as we might have wished. In the thirds space, the rule is clearer. Namely, chords are always built up expanding in an outward sense, treating the two axis (major and minor third) impartially (barring optional stylistic

adjustments in the case of ninth chords, to avoid discords). The diatonic scale is such that from any degree in the diatonic scale, one of these choices is always available and one always blocked if we wish to remain in the scale. This rule accounts, for example for the shape of the diminished triad on VII (Fig 8). See fig 9 for the appearance of scale-tone sevenths in the thirds space.

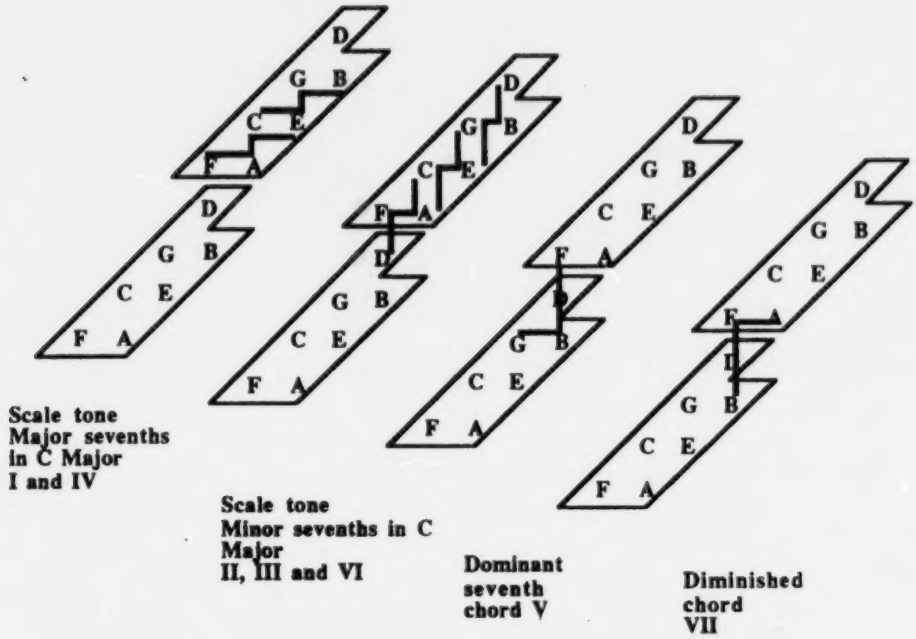


Fig 9 Scale-tone sevenths in the thirds space

4.2 Representing the 'Dynamics' of harmony

We can use the thirds version space version of Harmony Space to map out and control the possible movements of chord progressions. By moving along the cycle of fifths axis, progressions in fourths and fifths can be controlled. Trajectories of different lengths down this axis are shown in Fig 10. Depending on whether notes outside the key window are allowed or disallowed as roots of chords, real or diatonic progressions result. By restricting the chord vocabulary so that only some notes of the diatonic scale are available to be sounded, many basic and important chord progressions can be played as straight line gestures (e.g. I IV V I). When roots lying outside the key window are disallowed as roots, straight line gestures along the chromatic axis produce scalic progressions (fig 11).

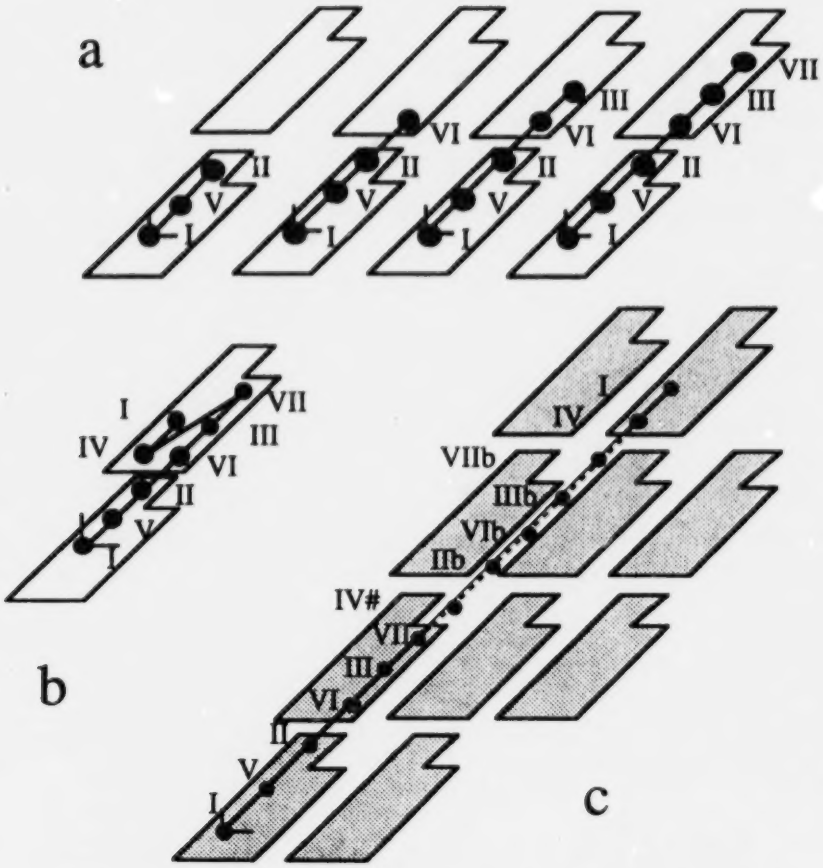


Fig 10 Trajectories down the fifths axis in the thirds space

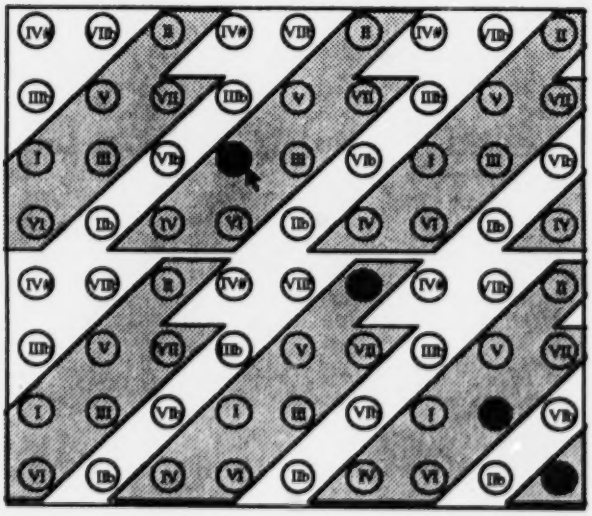


Fig 11 Scalic trajectories

When the key constraint is relaxed, chromatic progressions can be played on this axis (fig 12). As in the thirds/fifths version of Harmony Space, arbitrary chord progressions can be played, with discontinuous jumps being called for only in exceptional cases involving non-diatonic tones and tritones.

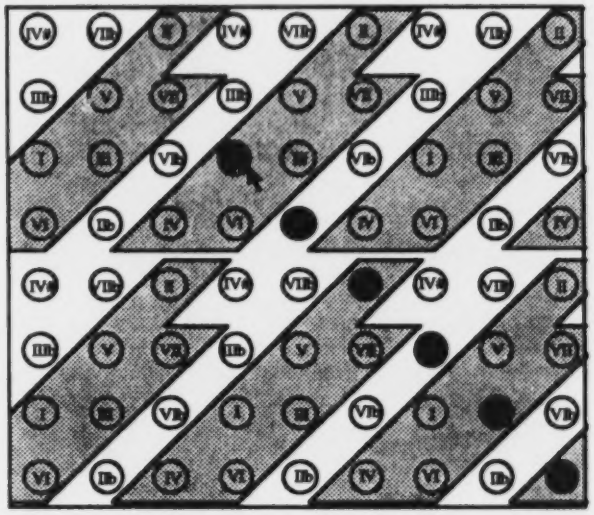
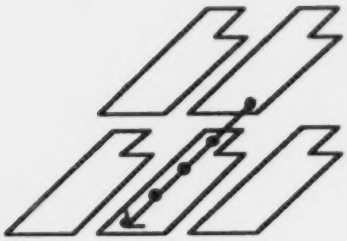


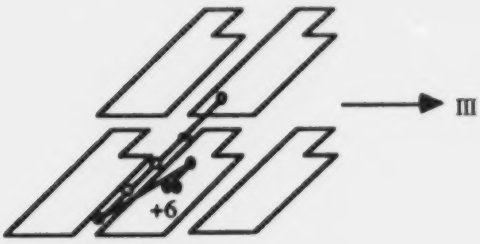
Fig 11 Chromatic trajectories

5 Analysing a chord progression in the thirds space

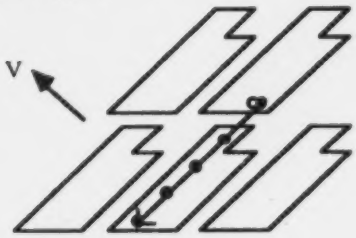
Fig 13 shows how the trace of "All the things you are" appears in the representation of the thirds version of Harmony Space. For clarity, only the root of each chord is indicated, and chord alteration is indicated by annotation.



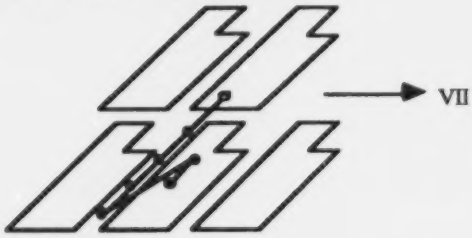
bars 1-5
Starting key I



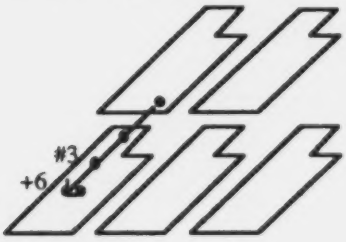
bars 6-8
Modulates up a major third in bar 6



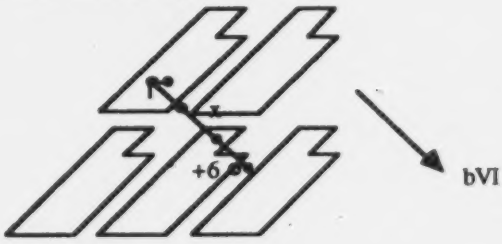
bars 9-13
Modulates up a minor third in bar 9



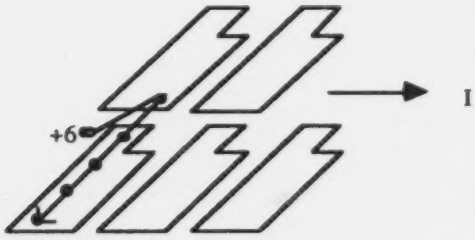
bars 14-15
Modulates up major third in bar 14



bars 16-20

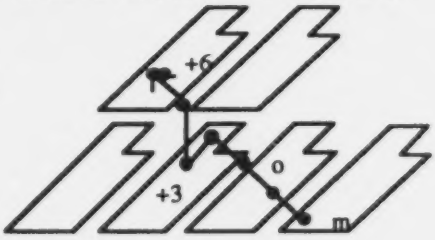


bars 21 - 24
Modulates down a minor third in bar 21



bars 24-28
Modulates up a major third in bar 24

(Arrows show direction of key window shift)



bars 29-36

(Ab) VI / II / V / I / IV / (C) V / I / I+6 /
(Eb) VI / II / V / I / IV / (G) V / I / VI / II /
(G) V#3 / I / I+6 / (E) II / bIIx / I / I+6 /
(Ab) VI / II / V / I / IV / IVm / III / bIIIo / II /
(Ab) V#3 bIIx / I+6 / I+6 //

Fig 13 Trace of "All the Things You Are" in the thirds space

(Chord sequence reproduced from Mehegan (1959) "All the things you are" by Kern-Hammerstein © Chappell & Co., Inc. and T.B. Harms.)

6 Relationship between alternative Harmony Space designs

Although the Balzano thirds space and Longuet-Higgins 12-fold space have such different starting points, the resulting co-ordinate spaces are related by a simple shear operation. (Indeed, with suitable metrics, the two spaces can be treated as identical.) This means that the exactly the same design for Harmony Space can be implemented in either configuration with no differences at all bar a sheared screen display. This is not to say that the two theories are *psychologically* indistinguishable - indeed, psychological experiments have been carried out to try to find out which is the more accurate cognitive model by exploiting their underlying differences (Watkins and Dyson, 1985).[3] Neither does it follow that interfaces based on the two configurations would necessarily be *educationally* indistinguishable: some phenomena appear to be more perspicacious in one version of Harmony Space than the other (as we will investigate in a moment). Neither is it by any means a foregone conclusions that all of the 'straight line' axes in the Longuet-Higgins 12-fold space will be mapped onto some equivalent axis in the sheared space. This is easy to demonstrate by a counter-example. For example, if the Longuet-Higgins space were to be sheared vertically upwards, the chromatic axis would be stretched out of existence.[4] The following table shows how the four intervals that can form a well-defined co-ordinate system for the chromatic scale (perfect fifth, semitone, major third and minor third) are remapped into points of the compass if the Longuet-Higgins 12-fold space is sheared horizontally and vertically respectively.

Interval	LH 12-fold	horizontal shear	vertical shear
p5	N	NE	N
M3	E	E	NE
m3	SE	N	E
m2	NE	SE	Lost: SE axis becomes a whole-tone axis

[3] It appears to have been concluded that both theories have complementary roles to play as cognitive models.

[4] The 'northern' axis would be unaffected and the 'eastern' and 'south eastern' axes would be shifted anticlockwise through 45°. But the 'north-eastern' chromatic axis would be "stretched out of existence" northwards, and a whole-tone axis "squeezed" into existence from the south. See the table above for another way of representing the same thing.

The vertical shear illustrates the result of a shear operation in the general case, which is that proximity relations are destroyed on one diagonal axis. A sideways shear on the Longuet-Higgins 12-fold space is an exception to this general rule, since the 12-fold space repeats itself every three elements along the horizontal axis (chapter 5 fig 2). Due to this periodicity of three, when the Longuet-Higgins space is sheared sideways, the old chromatic axis is "stretched out of existence" but a new chromatic axis is sheared into being from the opposite direction. As a result of this state of affairs, all of the axes present in the Longuet-Higgins 12-fold space are preserved in the thirds space in some form.

This means that all results in the thirds space are transformed versions of patterns previously explored in Longuet-Higgins 12-fold space. Despite the near equivalence of the two interfaces from such a point of view, they are well worth investigating separately, for at least three reasons. Firstly, the introduction of a new theoretical rationale for the interface has allowed us to review our design decisions and suggest possible theoretically motivated variations on the design. Secondly, some relationships may be visually more perspicacious in one representation than the other. Thirdly and finally, the different theoretical underpinnings of the two interfaces tend to make one or the other better suited for teaching particular aspects of music theory. For example, a comparison of the structure of the pentatonic scale, Guido's hexachord and the diatonic scale could probably be carried out more naturally in terms of Balzano's theory.

We will now collect together phenomena that appear to be represented more perspicaciously in one version of Harmony Space than the other.

- The "chord growing" algorithm is visually clearer in Balzano's space - the chord always grows outward along one or other axis.

- In the Longuet-Higgins 12-fold space, there are four maximally compact three element shapes that fit into the key window, only two of which are triads (see chapter 5). In the thirds space, the major and minor triads are the *only* maximally compact three element shapes that fit into the key window. This provides a useful story to tell novices about the unique importance of

triads.

• The next point only holds in one variant of Harmony Space. In this variant, the notes along each axis ascend in the specified interval steps as normal, but *without* being mapped into a single octave - the notes just carry on ascending. In this variant tool, it turns out that the Balzano representation is better suited for playing melodies than the Longuet-Higgins version, for the following reason. In the Balzano version, the chromatic diagonal really is, (and sounds like) a semitonal axis, since it corresponds to the following interval; 'a major third - a minor third = a semitone'. On the other hand, using the Longuet-Higgins representation, the chromatic axis corresponds to and sounds like the following interval; 'a major third + a perfect fifth = a major seventh'. Hence in this particular variant of Harmony Space, the Balzano representation is a much more natural instrument for integrating the exploration of harmony and melody.

• We speculate that the Longuet-Higgins 12-fold representation may have one very important general advantage. It may be easier for novices to remember and reproduce the key window shapes, chord sequences traces and chord shapes based on the Longuet-Higgins 12-fold space, due to their relatively straightforward rectilinear shapes, as opposed to the perhaps more deceptive slanting shapes of the thirds space. This speculation remains open to experimental testing.

We might now ask which of the two representations gives rise to the 'best' version of Harmony Space. From an abstract point of view, if we consider the space we are interested in to be a 12-note pitch set (i.e. we are not interested in notational distinctions) then Balzano's is *the* canonical co-ordinate system for the space. We can easily see that the Longuet-Higgins co-ordinate system contains redundancy in this characterisation of pitch, because the fifths axis spans the space by itself. Both interfaces probably have unique contributions to make in different educational contexts depending on the characterisation of pitch we wish to use. However, experimental testing would be required to discover which version is educationally most successful; effects such as the different degree of memorability of different

shapes might turn out to have a stronger bearing than such abstract issues.

7 Implementation of Harmony Space

A prototype version of Harmony Space has been implemented with the Balzano screen configuration. The prototype exhibits, in some form, all of the key design points described in the minimum specification of Harmony Space (given in chapter 5). Some of the minor features are either not implemented or have bugs associated with them. A full description of the prototype and its limitations is given in appendix 1. A sample screen dump from the prototype is shown in fig 14.

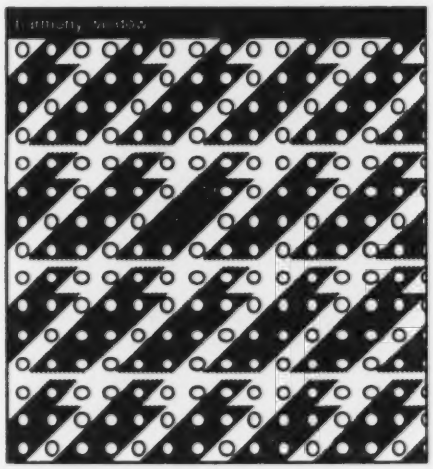


Fig 14 Screen dump from Harmony Space prototype[5].

The prototype implementation lacks proper menus and labelling, runs slowly and pauses frequently for garbage collection. The addition of a small number of features such as labels and menus would undoubtedly make it very much more convenient for novices. The prototype is clumsy to use because of these limitations, but is adequate to demonstrate, in rudimentary form, the educational potential of the underlying design. A report on some experiences of novices with Harmony Space can be found in chapter 9. There we are more easily able to discuss ideas made use of in this chapter not introduced until Part III. In appendix 2, a specification for a less

[5] For technical reasons to do with the the difficulty of printing large expanses of black, we have presented all of the screen dumps of the prototype inverted black-for white. Since the actual screen dumps are over 1Mb each in size - too large to include in the thesis document - they have been redrawn manually.

rudimentary version of Harmony Space is given (referred to as the 'idealised' or notional version - able amongst other things to use the layout of either the Longuet-Higgins 12-note or Balzano theories). This idealised version of Harmony Space is only trivially different from the prototype in conceptual terms. Both the implemented and idealised versions conform to the minimal specification given in chapter 5.

8 Some educational uses of Harmony Space

Harmony Space has at least five classes of educational use;

- sketching chord sequences,
- experimentally modifying existing chord sequences,
- exploring strategies for accompaniment
- informally analysing chord sequences,
- learning about music theory.

Examples of each of these activities as carried out by novices are discussed in Chapter 9. 'Sketching' chord sequences encompasses two distinct activities: trying to reproduce chord sequences that are already known and making up new ones. 'Experimentally modifying' chord sequences involves taking existing sequences and modifying them in musically informed ways. For example: progressions that make [illegible]f long cycles of fifths can be modified making use of tritone substitution (see section 4.4.3); chord sequences that use a consistent chord-arity can have that arity experimentally changed; chords can be substituted for other chords, for example, VII for V or II for IV (experiments with dominant and subdominant function); notes of altered quality can be substituted for other altered qualities in different melodic contexts, and so forth). Each modification can be exploited educationally in at least three ways. Firstly, the student can exercise aural imagination by trying to anticipate the musical result. Secondly she may use the modifications to try find out by trial and error which aspects of the piece are essential, and which parts can be altered without destroying the musical sense. Thirdly, she may try to decide whether or not changes are improvements, and if she considers they are, she has begun to compose 'new' pieces in a rudimentary way. Harmony Space is well-suited for exploring simple harmonic accompaniment strategies in combination with the human voice or other instruments. In particular, many of the strategies suggested by Campbell (1982) for

exploring accompaniment become relatively easy to try out using Harmony Space. For example, accompaniments using I alone, I and V alone, or I IV and V in combination can be explored using a single bass voice, scale-tone fifths diads or conventional triads, depending on the chord-arity chosen. Harmony Space also allows melodies to be 'tracked' in diatonic thirds or sixths, by selecting a chord arity of two. Tracking in sixths is a fundamental accompaniment strategy used with great frequency by both classical and popular composers (for example, "Oh Lovely Peace", (G.F.Handel) and "I feel fine" (Lennon & McCartney). Typical jazz accompaniment techniques can be explored by setting the chord arity to four (scale-tone sevenths) or five (ninths). Scale-tone sevenths are fundamental to the accompaniment of much jazz and form its basic harmonic raw material (Mehegan (1985). Scale tone ninths (chord arity 5) can give an accompaniment a distinctive jazz texture (cf. "The Nightfly" (Donald Fagen) and "Isn't She lovely" (Stevie Wonder)).

Novices can "informally analyse chord sequences" in the sense that, by watching the playback of a chord sequence in Harmony Space (by a human demonstrator or hypothetical implementation with built-in sequencer), they can produce simple functional analyses of the harmony of the piece. (It may also be useful for novices to analyse chord sequences informally at a higher level, in which whole sections of the chord progression are grouped as single units, as in the analysis of "All the things you are". In this approach to analysis, the interface is used as a notation which makes changes in direction of progressions, and their possible destinations graphically clear. These aspects of chord progression are not obvious in conventional notations). Finally, aspects of the theory of harmony can be explained and discussed using visual, audible and kinaesthetic demonstrations. Possible theoretical topics suitable for illustration to novices using version of Harmony Space include versions of practically every piece of music theory discussed in this chapter, for example: the nature of triads and tonal centres, the structure of the key system, modulation, the relationship of chord quality degree to degree of scale, making sense of chord progressions, simple modal harmony, the relationship between the major , minor modes and other modes and so forth. Note that as currently designed, it is expected that novices would require guidance in the use of Harmony Space - there is no claim that simple exposure to Harmony Space without guidance would by itself teach

novices to compose.

9 Intended users of the tool

The tool has been identified as being particularly relevant to the needs of those with no musical instrument skills, unfamiliar with music theory and unable to read music, but who would like to compose, analyse or articulate intuitions about chord sequences. Without such a tool, it is unclear how members of this group could even begin to perform these tasks without lengthy preliminary training. Thus the tool is not only an educational aid (though it can be used in this way) but an empowering tool that enables a group to perform tasks otherwise scarcely accessible at all. The interface can be used by novices to experiment with, compose, analyse, modify and play harmonic sequences in purposive ways. (Some of these tasks might be easier to carry out using modified versions of the interface - discussed in the section on conclusions and further work below.) Harmony space may be able to give conceptual tools for articulating intuitions about chord sequences not only to novices but also to experts. The graphic linear notation for chord progressions that emerges from Harmony Space when tracking chord sequences makes various aspects of chord progressions visible and explicit that are otherwise hard to articulate in conventional notations.

10 Related interfaces

To the best of our knowledge, the work described in this chapter was the first application of Longuet-Higgins' theory for educational purposes. Holland (1986b) appears to be the first discussion of use of Longuet-Higgins' representation for controlling a musical instrument, and the implemented prototype appears to be the first such instrument constructed. A number of interfaces, each related in some way to Harmony Space are described below. Each, including Harmony Space (Holland 1986b), was designed with no knowledge of any of the other interfaces. The first device is Longuet-Higgins' light organ (fig 15). Longuet-Higgins (reported in Steedman 1972 page 127) connected each key of an electronic organ to an square array of light bulbs illuminating note names. This device was the first to make explicit use of Longuet-Higgins' theory. It allowed music played on the organ to be displayed in Longuet-Higgins' original (non-repeating) space.

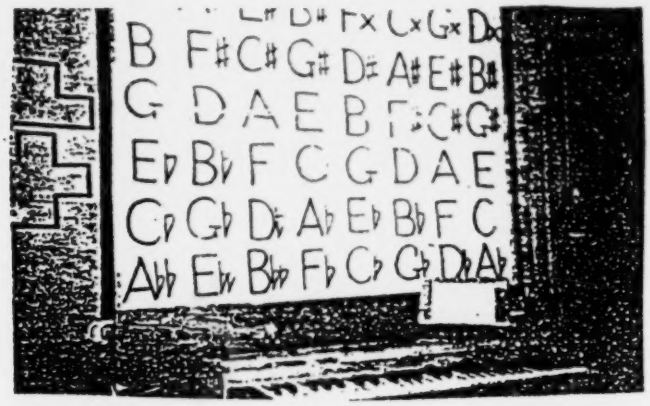


Fig 15 Longuet-Higgins' Light organ
(from Steedman, 1972 page 127)

However, the question did not arise of working in the opposite direction to allow the grid representation to control the organ. The key window does not seem to have been represented on the display. The first computer-controlled device using a generalised two-dimensional note-array (we will explain this in a moment) with pointing device to control a musical instrument seems to be Levitt's program Harmony Grid (Levitt and Joffe, 1986). Harmony Grid runs on an Apple Macintosh. It displays a two dimensional grid of notes where the interval between adjacent notes may be adjusted independently for the x and y axes to any arbitrary number of semitone steps (fig 16). The grid display can control or display the output of any musical instrument with a MIDI interface (Musical Instrument Digital Interface - an industry interconnection standard for musical instruments). Harmony Grid can be configured as a special case to the grid layouts of Balzano's space or Longuet-Higgins 12 fold space. The question of key windows does not arise in Harmony Grid. This means that Harmony Grid does not make explicit the bulk of the relationships and structures described in this chapter. The mouse can control chords, but their quality must be adjusted manually - there is no notion of inheriting or constraining chord quality from a key window. Hence modulation can be carried out only by knowing and manually selecting the appropriate chord quality as each chord is played. Harmony Grid is the first of its kind and a useful educational tool. It is robustly implemented with a good real-time response and has several useful performance-oriented features.

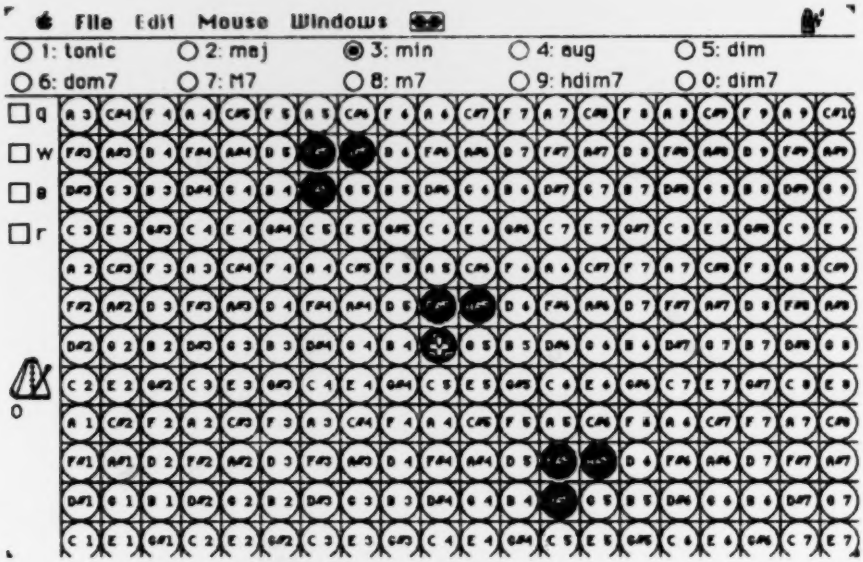


Fig 16 Harmony Grid

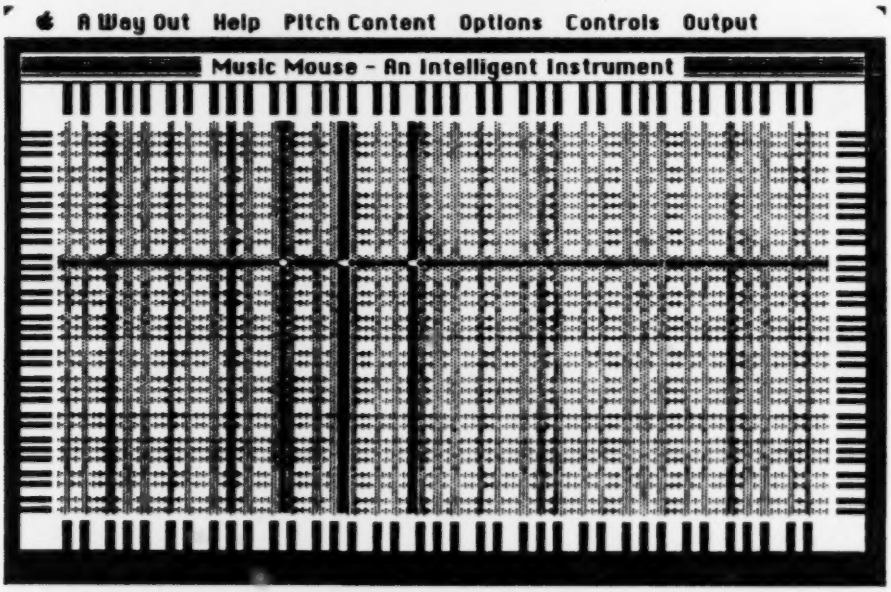


Fig 17 Music Mouse

Another related program is Laurie Spiegel's (1986) "Music Mouse" (fig 17). Music mouse has two piano-type keyboards represented at right angles on a computer screen. Three vertical bars and one horizontal bar are displayed that extend over the height and width of the screen respectively. The horizontal bar sounds melody notes as it is moved up and down, brushing the vertical keyboard. The vertical bars sound

chords as they are moved forward and backwards over the horizontal keyboard. The vertical bars move together, the horizontal position of the group controlled by the x position of the mouse. These sound an inverted triad (or another selectable chord). Simultaneously, the y-position of the mouse controls the position of the single melody note represented by the horizontal bar. All four notes are constrained by the system to avoid black notes and play only white notes (other constraints are selectable). Hence key constraints can be enforced. Music Mouse is a highly original program, and a delightful and sophisticated instrument to use. Not surprisingly, it does not make harmonic relationships anything like as explicit as Harmony Space does. It is easy to play linear chord progressions, but hard to control or recognise progressions in mixtures of intervals and there is no way of visualising inter-key relationships. Harmony Space is by far the most expressive of these interfaces from the point of view of of explicit control and representation of harmonic relationships. Balzano (1987) has recently been working on the design of computer-based educational tools for learning about music. One of the tools is based on his group-theoretic approach to harmony, with a two-dimensional visual display (Balzano, 1987), but no details appear to be available in the literature yet.

Finally, a development of Harmony Space called *Harmony Logo* has been developed, in prototype, as part of this research. Harmony Logo is a programming language that moves a voice audibly around in Harmony Space under program control. It does this in a fashion analogous to the way the programming language "Turtle Logo" moves graphic turtles around on a computer screen. The voice or homophonic instrument is, by analogy, a "harmony turtle". Harmony Space has the potential to allow the structure of long bass lines and chord sequences to be compactly represented in cases where the gestural representation might be awkward to perform or hard to analyse. Harmony Logo plays a role in the architecture of Chapter 8.

11 Limitations

The current implementation of Harmony Space is an experimental prototype designed to show the feasibility and coherence of the design. At the time at which the work was done, the graphics, mouse and midi programming in themselves required

considerable effort to undertake. The prototype served its intended purpose, but is too slow and lacks simple graphic and control refinements to make it easy to use. In particular, it suffers from the following limitations;

- the prototype is slow to respond and garbage collects frequently, hampering rhythmic aspects of playing the interface,
- the slow response of the prototype means that events that should emerge as single gestalts (such as smooth chord progressions and modulations) are broken up,
- the response problems interfere with the illusion of a direct connection between gesture and response, and can cause considerable frustration,
- the secondary functions such as overiding chord quality, switching chord arity, controlling diatonicism/chromaticism and modulating are controlled in the current prototype by function keys as opposed to pointing devices,
- the 'state' that the interface is in, and the choice of states available is invisible, rather than indicated by buttons and menus on the screen,
- it is not possible to view the grid labelled by roman numeral, degrees of the scale, note names etc.,
- output is not provided in Common music notation, 'piano keyboard and dot' animation or roman numeral notation to assist skill transfer (see appendix 2).

Other limitations of Harmony Space are described in detail in Chapter 10.

12 Conclusions

The following conclusions all apply to *both* versions of Harmony Space we have presented (the Longuet-Higgins 12-fold version and thirds space version), except for those otherwise indicated.

12.1 Design decisions and expressivity

A theoretical basis and a detailed design for a computer-based tool for exploring tonal harmony have been presented. The design exploits recent cognitive theories of how people perceive and process tonal harmony. Design decisions have been exhibited which lead the representations underlying the theories to be incorporated into the

interface in ways which give the interface particularly desirable properties. The design decisions and the properties that they give rise to are summarised below. Firstly, the relevant design decisions are as follows;

- use of the 12-fold (as opposed to non-repeating) space[6],
- replication of the 12-fold field and key window graphically to allow linear patterns to emerge,
- use of the entire replicated pattern as a direct manipulation control and display area,
- use of the same spatial metaphor for controlling and displaying notes, chords and modulation,
- provision of scale-tone chord qualities by default,
- visual metaphor for scale-tone default,
- separation of control of chord-arity from control of chord quality,
- capability to overide default scale tone qualities,
- synchronisation of three modalities (auditory, haptic and visual) for control and representation,
- optional control of diatonicity (control of access to chromatic notes as roots of chords),
- use of more than one pointing device.

These design decisions and the underlying cognitive theories give the resulting interface the properties listed below. In some cases, three of the most elegant and ingenious traditional 'interfaces' are used as points of comparison: common music notation, the guitar fretboard and the piano keyboard.

1 Chord progressions fundamental to western tonal harmony (such as cycles of fifths and scalic progressions) are mapped into simple straight line movements along different axes. This makes fundamental chord progressions easy to recognise and control in any key using very simple gestures. (These progressions can be difficult for novices to control and recognise by conventional means.)

2 The same property allows particularly important progressions (such as

[6] Applies only to the version with the Longuet-Higgins' configuration.

cycles of fifths) to be controlled and recognised as *unitary "chunks"*. A common problem for novice musicians using traditional instruments or notations is that they only deal with chord progressions chord by chord (Cork, 1988).

3 Chords that are harmonically *close* are rendered visually and haptically close. This makes other fundamental chord progressions (such as IV V I and progression in thirds) easy to recognise and control gesturally. (On a piano, for example, these progressions involve gestural 'leaps' that can be particularly hard to find in 'remote' keys)

4 Similarly, *keys* that are harmonically close are rendered visually and haptically close. This makes common modulations easy to recognise and control. (In common music notation, or on either traditional instrument , considerable experience or knowledge can be required to abstractly calculate key relationships.)

5 The basic materials of tonal harmony (triads) are rendered visually recognisable as maximally compact objects with three distinguishable elements. (In other representations, the selection of triads as basic materials can appear arbitrary.)

6 Major and minor chords are rendered easy to distinguish visually. (This is not the case, for example, in common music notation or on a piano keyboard)

7 Scale-tone chord construction is rendered such that it can be seen to follow a consistent and relatively simple graphic rule in any key. (On a piano keyboard, for example, this is not true in remote keys, where semitone counting or memorisation of the diatonic pattern for the key is required)

8 Consistent *chord qualities* can seen to be mapped into consistent shapes, irrespective of key context. (Again, this is not true, for example, on a piano)

9 The tonic chords of the major and minor keys are seen as spatially central within the set of diatonic notes, whatever the key context. (This relationship is not explicit in common music notation, guitar or piano keyboard)

10 Scale-tone chords appropriate to each degree of the scale can be played (even when there is frequent transient modulation) with no cognitive or manipulative load to select appropriate chord qualities (This is not the case with guitar or piano).

11 The rule governing the correspondence of chord quality with degree of scale can be communicated by a simple, consistent physical containment metaphor.

12 A single consistent spatial metaphor is used to describe interval, chord progression, degree of the scale, and modulation. (Not the case in the other three points of comparison)

12.2 Uses of Harmony Space

The following uses for Harmony Space and its possible variants have been identified;

- musical instrument,
- tool for musical analysis,
- learning tool for simple tonal music theory,
- learning tool for exploring more advanced aspects of harmony, (e.g. assessing the harmonic resources of other scales)
- discovery learning tool for composing chord sequences,
- notation for aspects of chord sequences not obvious in conventional notations.

12.3 Longuet-Higgins' theory and education

The report from which this part of the thesis stems (Holland, 1986b) appears to be the first discussion of educational use of Longuet-Higgins' theory (whether by

electronic or other methods), and its use for controlling as well as representing music.

Fuller conclusions about Harmony Space are given in the final chapter of the thesis.

Part III: Towards an intelligent tutor for music composition

Chapter 8: An architecture for a knowledge-based tutor for music composition

"You must explain what they can do.
They don't understand so much freedom"
Harpe (1976, Page 4)

1 Introduction

In this chapter, we present a framework for a knowledge-based tutoring system to help novices explore musical ideas in Harmony Space. This framework is referred to as MC. This framework was presented at the Third International Conference on AI and Education (Holland, 1987d), with an extended abstract in Holland (1987c), and at the First Symposium on Music and the Cognitive Sciences (Holland, 1988). Aspects of it are also discussed in Holland and Elsom-Cook (in press). The framework is aimed particularly at providing support for composing chord sequences and bass lines.

2 Tutoring in open-ended creative domains

The motivation behind the framework can perhaps most clearly expressed by starting from elements of a working definition of creativity proposed by Johnson-Laird (1988). In many respects Johnson-Laird's view of creativity is similar to ideas expressed by Levitt (1981), Levitt (1985), Sloboda (1985) and Minsky (1981), though not stated in these earlier contexts quite so generally. For our purposes, we need only consider two elements from Johnson-Laird's definition. The first element of the definition is the assumption that creative tasks cannot proceed from nothing: that some initial building blocks are required. The second element is the assumption that a hall-mark of a creative task is that there is no precise goal, but only some pre-existing constraints or criteria that must be met (Johnson-Laird,1988). From this

starting point, the act of creation can be characterised in terms of the iterative posing and eventual satisfaction (or weakening or abandonment) of constraints by the artist. The artist may at any time add *new* constraints to the weak starting criteria.

Sometimes, and in some domains, it may be acceptable to sacrifice a pre-existing constraint or criterion in order to meet new constraints imposed by the artist: Sloboda (1985) puts this very clearly;

> " ..we will find composers breaking .. rules [specifying the permissable compositional options] from time to time when they consider some other organisational principle to take precedence." (Sloboda, 1985, Page 55)

We will refer to creative domains in which pre-existing constraints and criteria may get broken in this way as "open-ended creative domains".

It is scarcely new to describe artistic creation in terms of imposing, loosening and satisfying constraints of various sorts; perceptual, cultural, stylistic and personal. Artists have been using such descriptions informally for a long time. But most formal theories of music are framed according to a different outlook: one based on grammars and rules. This choice may have subtly misled theorists and educators: grammars and rules are not normally geared for the dynamic devising and weakening or abandonment of rules in mid-application (though see, for example, Lenat (1982) and Holland (1984b)). A grammatical perspective may tend to encourage us to think of each layer of constraints as existing as a more or less closed layer, devaluing the importance of interactions with the other levels. Worse yet, it may encourage us to think of some layers, such as the 'rules' of four part harmony as being especially sacrosanct, leaving us with no simple way to account for violations that inevitably occur. In reality, it appears that a composer may sacrifice constraints at *any* level if she feels that constraints at other levels are more important in a particular situation.

2.1 The nature of constraints in music

Constraints in music seem to fall broadly into three types: some based on fostering perceptual and cognitive conditions for effective communication, others based on

cultural consensus and yet others introduced *de novo* by the artist. We will look at this distinction in a little more detail, as it has been a major influence on the architecture of MC. Examples of the first kind of constraint appear in recent research on western tonal music, such as work by Balzano (1980), Minsky (1981), McAdams and Bregman (1985), concerned respectively with harmony, metre and voice leading. Each of these pieces of research emphasises how various widespread features of music appear to have an important role in fostering perceptual and cognitive conditions for effective communication of structure. (Work of this sort, such as Lerdahl and Jackendoff (1983) is typically couched in grammatical terms rather than in terms of constraint satisfaction, although there is no particular difficulty in re-expressing such considerations in terms of constraints.)

The second class of constraints is more of a cultural, historical nature. The nature of music is not only affected by the structure of musical materials and how these interact with our perceptual and cognitive faculties, but also by listeners' familiarity with the way in which composers happen to have used these materials to date. The particular genres that happen to have evolved and the particular pieces of music that happen to have been written necessarily have powerful effects on modern listeners and composers. When listeners hear a new piece of music, cognitive theories of listening posit that the music must be chunked in various ways to cope with memory and processing limitations (Sloboda, 1985). The kind of chunking that can be done by a particular listener depends not only on the abstract nature of materials such as harmony, but on the practices, genres and particular pieces of music with which the listener is already familiar. Levitt puts the connection between stylistic constraints and those introduced by an individual composer very well (Levitt, 1981 2.2.1);

> " Effective communication requires musicians to repeat structures frequently within a piece and collectively over many pieces. Usually we view "musical style" and "theme and variation" as utterly different. Computationally and socially they are similar things with different time spans; style tries to exploit long term "cultural memory", while theme and variation exploits recent events. In either case, the considerate composer uses an idea of what is already in the audiences head to make the piece understandable."
>
> (Levitt, 1981 p. 15)

2.2 *Constraints and pedagogy*

The characterisation of creative tasks in terms of constraints fits well with commonplace intuitions and practices in creative domains. A ubiquitous technique in the arts to make a creative task easier is to *impose more constraints* - sometimes almost irrespective of the exact nature of the constraints. Many educators in the creative arts, such as Bill and Wendy Harpe (Harpe, 1976) use this thematic technique routinely and explicitly, for example, allowing mural painters to paint anything so long as it is entirely in shades of blue (Harpe, 1976) or commissioning a musique concrète composer to work with a free hand as long as no sounds were used except building site sounds (Great Georges Project, 1980). Paradoxically, approaching a highly constrained creative task can be subjectively much easier than approaching a tabula rasa.

Johnson-Laird's characterisation illuminates pitfalls in the standard way of teaching four part harmony. The text book rules for teaching traditional four part harmony have been abstracted from the historical practice of composers. These 'rules' are often treated as if they were privileged compared with any constraints that the composer herself might bring to the exercise. This may be an effective pedagogical device, but one with damaging side effects if it leads the learner to believe that composing music is mainly a matter of conforming to rules derived from previous practice. As we have noted, composers appear to work with many different kinds of constraints, some derived from previous practice, but many specific to the piece in hand. When it becomes impossible to satisfy all of the constraints, those derived from previous practice may have no particularly privileged status.

The foregoing has implications for the construction of tutoring systems for creative domains. When constructing an intelligent tutor for four part harmony or 16th Century counterpoint, it may be justified to follow standard pedagogical practice and adopt the fiction that there is some certified knowledge about the domain that can be applied more or less rigidly in a prescriptive and critical fashion. However, in the case of still evolving popular music, this fiction is untenable. The recent history of popular music is littered with examples of naive musicians who have written successful pieces of popular music that obey internal constraints of their own yet depart markedly from existing practice. A remedial tutor in this domain is

inappropriate since its prescriptions run the risk of being misguided and its criticisms of being misconceived.

However, this does not mean that a student is better left with no guidance at all. As Johnson-Laird's characterisation emphasises, initial building blocks are essential. Furthermore, informal observation of novices learning in informal situation suggests that many kinds of information can help a novice to learn to compose, for example;

- pieces of music that exemplify a prototypical use of some musical material,
- pieces of music that exemplify a prototypical use of some strategic idea,
- ways of making a piece sound as if it is in a certain style,
- ways of employing various musical devices or techniques
- ways of employing various musical strategic ideas,
- information about how a particular piece gets a particular effect.

In general terms, the idea behind the MC architecture is to work at three levels to be illustrated shortly - song level, style level and musical plan level. Where possible, we represent musical knowledge in terms of constraints, so as to fit with notions of creativity as outlined above. However, in the case of the lowest level of representation, we work with Harmony Space-based representations, so as to ease the burden of communicating harmonic concepts to novices. An interlinked tutor and microworld is used in a way that fits well with the idea of Guided Discovery Learning (Elsom-Cook, in press). We will begin by outlining the broad architecture of the framework, and an example interaction style, before looking at each representation level in turn.

3 Architecture of a knowledge-based tutor for music composition

In broad terms, as we indicated in the introduction, our framework for providing guidance in the exploration of the learning environment involves linking a knowledge-based tutor to one or more musical microworlds (Fig 1). The harmony microworld in fig 1 is Harmony Space. Integrated microworlds for rhythm and melody are not presented in the thesis. Such microworlds would clearly be desirable. In the further work section of the thesis as a whole we will identify possible theoretical bases for such microworlds and outline the form that they might

take. Fortunately, for our immediate purposes, a single microworld is sufficient to illustrate the issues to be discussed.

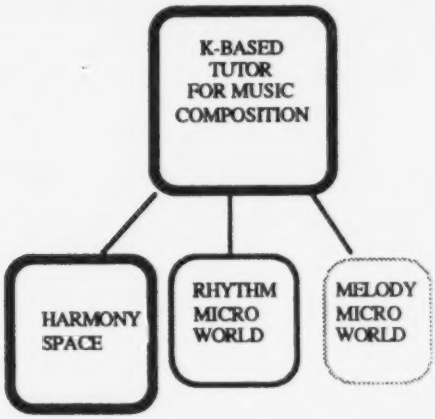


Fig 1 General architecture of proposed tutoring system

One aim of the architecture is that the knowledge base should provide abstract knowledge for the more effective exploration of the microworld: the microworld should provide experience in several modalities; visual, gestural and auditory, in order to make this abstract knowledge concrete in a principled way. This kind of approach may be viewed as a special case of Guided Discovery Learning, proposed by Elsom-Cook (1984) as a framework for combining the benefits of learning environments with those of interventionist tutors. Guided Discovery Learning methodology proposes the framework of a microworld in which discovery learning can take place, integrated with a tutor that is capable of executing teaching strategies that range from the almost completely non-interventionist to highly regimented strategies that give the tutor a much larger degree of control than the student.

Interventionist teaching strategies for providing individually relevant guidance are not presented in the thesis. The MC framework looks very well suited to supporting a wide range of active teaching strategies; for example, versions of all of those discussed by Spensley and Elsom-Cook (1988). However, the work on MC concentrated on finding solutions for the core problems that had to be solved to make a tutor possible at all: the devising of ways of characterising strategic considerations about chord sequences and the devising of formalisms for the song, style and plan levels that were potentially suitable for communication with novices. Prototypes of the song and plan levels have been implemented, and various musical

plans have been devised and encoded. (See later in this chapter for an implementation summary of MC) A very simple user-led strategy which the student could use to navigate through a series of transformations in the knowledge bases is presented. Teaching strategies to provide individually relevant guidance in MC are not discussed except in the further work section of this chapter.

3.1 An interaction style: the 'Paul Simon' method

Since we have decided that a prescriptive, critical tutoring style is inappropriate for the domain under consideration, we need at least one non-prescriptive, idiom-independent, general-purpose strategy we can offer novices for composing music. One simple and time-honoured candidate approach to composing music is the 'Paul Simon', or 'intelligent transformation' method. The Paul Simon method takes its name because according to Stephen Citron (1985 page 181) it is one of Simon's present day methods of composition. Very simply, the method involves taking an existing song that is liked, and then modifying the song successively. At a certain point, the result is no longer the original song, and can be claimed as an 'original' song. However, the method is not as simple as it looks. It would be quite inadequate for a novice to make piecemeal trial and error modifications blindly at note level. Without appropriate musical knowledge the hit rate of musically sensible modifications would be low. Most novices lack the musical knowledge to make changes that are predominantly musically sensible. The method can work for experienced musicians because they have extensive knowledge about what changes are musically sensible, but it is inherently limited. An accomplished composer will have far more powerful and flexible approaches at her command than adaption. However, it seems reasonable to assume that a novice who began by using the Paul Simon method and gained enough musical knowledge to apply the method well would have gained enough knowledge about musical structures, processes and genres to be in a position to develop her own compositional methods. It is not being suggested that this method is a particularly desirable composition method, nor does the approach of this chapter demand its use. All that is being claimed for now is that it is one example of a non-prescriptive strategy that can be used as a guideline to allow the user navigate herself through the knowledge base for the purpose of composing - any number of other strategies could be substituted. Later in the

chapter we will discuss further kinds of interaction that could be supported within the same framework.

What a knowledge-based tutor requires to support the Paul Simon or "intelligent transformation" method is some way of characterising what constitutes a musically sensible change. This is the question that we now need to address. The idea behind the architecture is to work at three levels - song level, style level and musical plan level (fig 2). We will now consider more closely what kinds of musical knowledge the architecture requires, taking each kind of knowledge in turn.

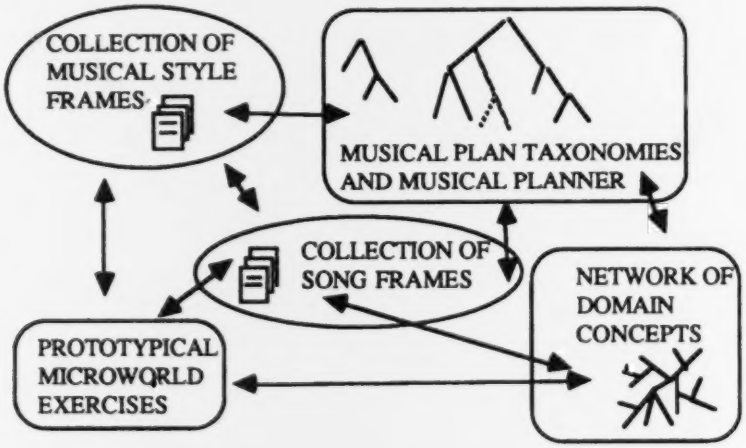


Fig 2 Kinds of domain knowledge in MC

3.2 Representation of Pieces

In music, case studies tend to have more weight than theory. Hence there is a clear need for an extensive net of examples to serve as a source of initial building blocks. The collection of case studies should include pieces likely to be known and liked by clients, and should include exemplars of a wide variety of musical styles, features, musical goals, techniques and devices. However, few novices could be expected to learn much by simply listening to or looking at pieces of music at note level. It may help if we can "chunk" music in various different ways and record interconnections between chunks.

What representation formalisms should be used to chunk descriptions of songs? Given our focus on chord sequences, Harmony Logo might be a good candidate as

a representation formalism at the 'song' or 'piece'[1] level of description. We will argue that descriptions in Harmony Logo can be used to chunk chord sequences in ways that can correspond well with musicians' intuitions. Harmony Logo has the additional pedagogical advantage that abstract, chunked descriptions of chord sequence structure framed in its terms could be given direct principled visual illustration in Harmony Space. The expositions of part II of the thesis illustrate the likelihood that commonly recurring harmonic concepts could be illustrated to novices by suitably designed practical exercises in Harmony Space. Other candidate formalisms for 'piece' or 'song' level descriptions could be constraint satisfaction representations similar to those developed in Levitt (1985), or hierarchical descriptions similar to those of Lerdahl and Jackendoff (1983). The latter two candidates are attractive in many respects, but appear to be less well adapted than Harmony Space to communicating basic harmonic concepts to novices (see Lerdahl and Potard (1985) and Levitt (1985)). Conceivably, several different formalisms could be used to complement each other, in line with Minsky's suggestions about the use of multiple complementary alternative descriptions of complex phenomena (Minsky, 1981b), and the increasing recognition of the need for multiple representations in Intelligent Tutoring Systems research (Stevens, Collins and Goldin, 1982; Self, 1988; Moyse, 1989).

Leaving aside the way in which they are represented formally, one useful conceptual model for such chunks is the kind of terms some musicians use informally with one another when discussing pieces semi-technically. Musicians are often able to communicate useful information concisely about a piece of music using a few "chunks" of information.

Verse - repeated twice - (harmonic rhythm of one chord per bar.)

A min7 D min7 E min7 A min7

Chorus (harmonic rhythm of two chords per bar.)

D min7 G7 C ma7 F ma7 B dim7 Edom7 A min7

Fig 3 Chord sequence of Randy Crawford's (1978) "Street Life"

[1] We will use the terms 'piece' and 'song' synonymously in Part III.

For example, to quickly communicate the essentials about the chord sequence of Randy Crawford's (1978) "Street Life" (Fig 3) - a musician might say;

> "The verse is a three chord trick, but in an aeolian version, using jazz sevenths". Similarly, for the chorus a musician might continue; "The chorus goes from the tonic right round the chords in a diatonic cycle of fifths, with the cadence in the harmonic minor".[2]

We will see in a moment that this kind of chunking is particularly useful in suggesting meaningful ways in which a song can be modified. Let us look at a slightly more detailed example of this kind of chunking, using a version of "Abracadabra" by the Steve Miller Band (1984), given in standard notation below[3];

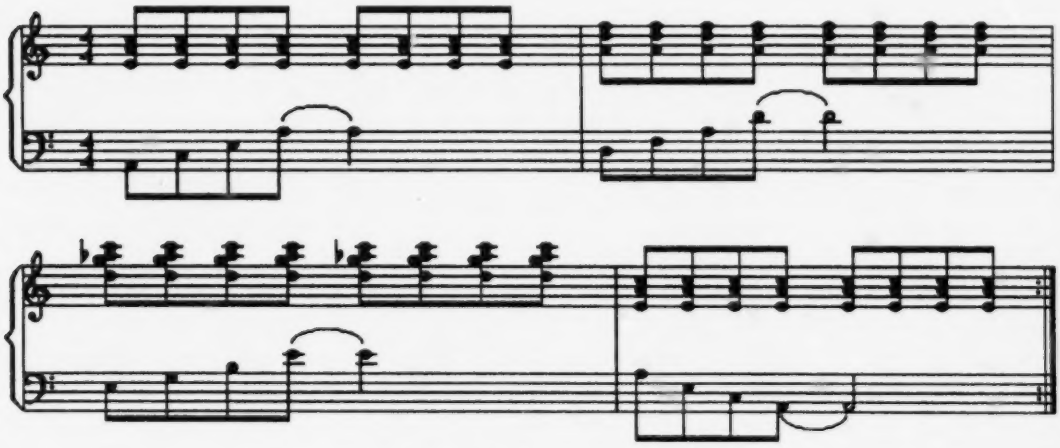


Fig 4

Bass line and chords for a version of

The Steve Miller Band's "Abracadabra" (Steve Miller, 1984)

We can encapsulate knowledge about this piece of music as follows. (We do not claim that this represents the complete or only true picture of this piece, but we will argue that it does give one useful view):

[2] Some readers may consider this kind of description arcane, but recently when trying to establish whether copyright clearance was required for use of musical examples in a version of this chapter being prepared for publication, I found myself describing the extracts over the telephone to the university copyrights officer (a musician) using just such terms.

[3] The Ab in the third bar is strictly a G#, represented as shown for notational convenience.

- The chord sequence is a straight line down the 'fifths axis',
- The piece has a three chord vocabulary,
- The piece is in the minor mode,
- The piece uses plain triads,
- The dominant chord employs a variant jazz quality (a ma/mi 3rd +13th),
- There is a motor rhythm in the right hand accompaniment,
- The bass line is a four note triadic aeolian arpeggio. This is played once per bar with the arpeggio upward but reversed in the last bar.

Chunking would not be very useful unless the chunks could be usefully manipulated. A feature of the MC architecture is a facility to generate a playable fragment from any isolated chunk or collection of chunks. Broadly speaking, this involves having defaults available that can be over-ruled by more specific information when available - as in Levitt's (1981) default-based improvising system. With such a facility, a novice could get a very rapid idea of what particular chunks did, by resynthesizing the song through stepwise refinement using each chunk in turn. We will now illustrate in words the kinds of transformations involved.

3.2.1 Illustration of elementary chunk manipulation

Starting with no chunks at all, and asking for the system to use its defaults to play the song description so far gives the following result.

(major) I I I I

This is a rather dull piece of music. However, it accurately reflects the basic assumptions of the system, as we will now explain. The repeated major tonic chord reflects the built-in assumptions that we are dealing primarily with tonal chord sequences (as opposed to melodies, atonal chords, musique concrète etc.); that the basic stuff of tonal chord sequences is triads (as opposed to diads, sevenths, ninths, etc); that we are dealing with music in a metrical context (where events tend to occur with reference to regular pulses); that in the absence of any other information about a chord sequence, the 'safest' thing to do in some sense is to repeat the tonic (as

opposed to a chord based on any other degree of the scale); that the default tonal centre is the major centre (as opposed to the minor or a modal centre).

In order for verbal explanations of some sort to be available to novices exploring HMC, we would need to associate canned text and pointers to a semantic net with each chunk and with each setting in the vanilla (default) setting. This is not implemented. However, it appears well within the range of standard ITS techniques (Stevens and Collins (1977); Wenger,1987; Elsom-Cook (in press)).

Moving on from the system defaults, the first chunk of information associated with the piece is; 'The chord sequence is a straight line down the fifths axis'. Given this chunk[4] and no other information save the system defaults, the description generates the following;

(major) I IV VII III VI II V I .....

Let us now move onto the second chunk, which is as follows; 'The piece has a three chord vocabulary'. This indicates that the song uses only three functionally distinct chords, and further more that these chords are I, IV and V. The effect of this chunk on top of the first chunk and the system defaults would be both aurally very marked and easy to demonstrate graphically in Harmony Space; the resulting chord sequence is as follows;

(major) I IV V I .....

The third chunk; 'The piece is in the minor mode' indicates that the song uses the minor (not major) version of the three chord trick. The four chunks together produce a change that would be both audible and visible in a Harmony Space trace. The corresponding chord sequence is as follows;

(minor) I IV V I

[4] The use of Harmony Space concepts in some chunks does not give rise to any technical problems irrespective of the representation used for chunks, since Harmony Space concepts are expressable in any general purpose programming language capable of dealing with mathematical and logical terms.

The fourth chunk; 'The piece uses plain triads' merely explicitly confirms the default assumption that the song uses plain triads (and not sevenths or ninths etc). This would lead to no audible change, no change in a Harmony Space animation, and no change in the chord sequence generated. The fifth chunk; 'The dominant chord employs a variant jazz quality (a ma/mi 3rd +13th)' indicates that the chord quality of the chord with root V is altered to a particular jazz variant. The chunks collected so far now yield the following slightly modified chord sequence;

(minor) I IV V (ma/mi 3rd +13th) I

The sixth chunk; 'There is a motor rhythm in the right hand accompaniment' indicates that the three chords are delivered using a motor rhythm (Cooper, 1981 page 57), with the collective chunks now producing;

Fig 5 - "Abracadabra" under assembly - motor rhythm stage

Finally, the seventh chunk: 'the bass line is a four note triadic aéolian arpeggio'. This is played once at the beginning of each bar (at the same metrical rate as the motor rhythm) with the arpeggio upward, but reversed in the last bar' finally yields;

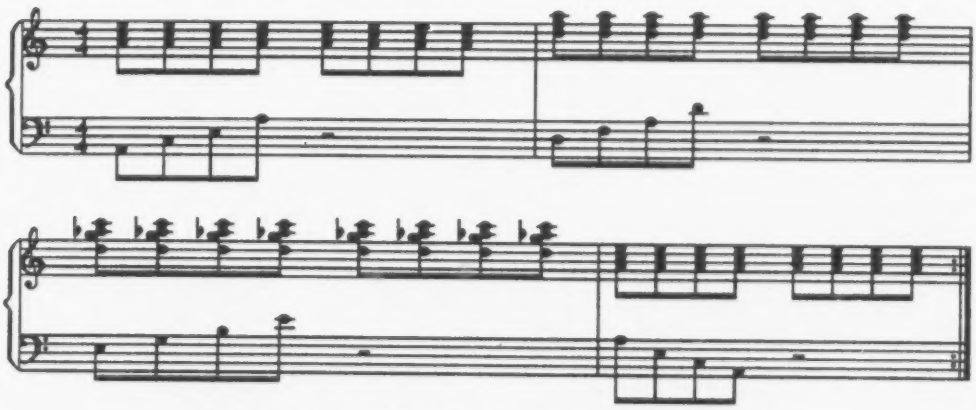


Fig 6 - "Abracadabra" reconstructed in seven chunks

In seven chunks, this has built up a reasonable approximation of the chord sequence and associated bass line of a version of "Abracadabra". Furthermore, in principle most of the chunks could be explained and illustrated using Harmony Space without any need for extensive prerequisite knowledge about music.

For a novice, the process of replacing each chunk in a chunked description of a piece with an appropriate alternative chunk has a number of possible benefits; the process provides a systematic opportunity to explore the chunks, explore the piece, and to search for musically sensible variations on the piece. In the present example, each chunk implicitly suggests a space of potential alternative chunks, although the size of the space of alternatives varies enormously among chunks and depends on the level of generality of class which a chunk is viewed as belonging to. For example, the minor 'mode' is a member of the set of modes, of which there are only a limited number. A choice of the major mode in place of the minor mode would yield the following;

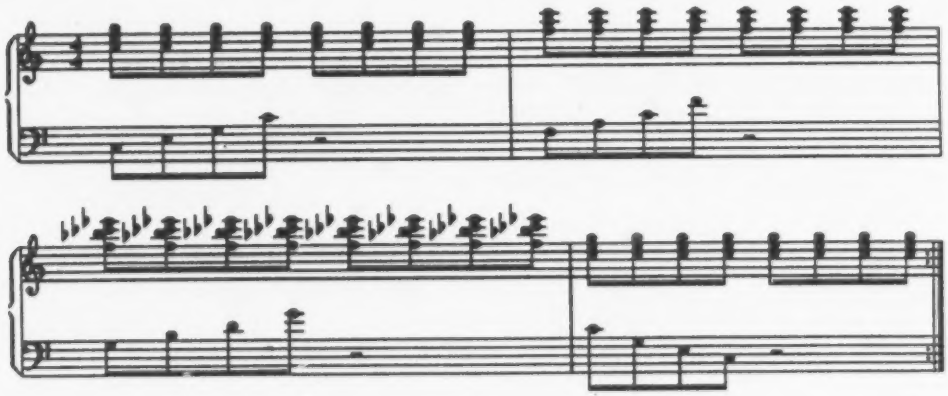


Fig 6a - Major version of "Abracadabra"

Similarly, the triad chord arity is one of a limited number of standard chord arities; a choice of ninth arity would cause the following changes;

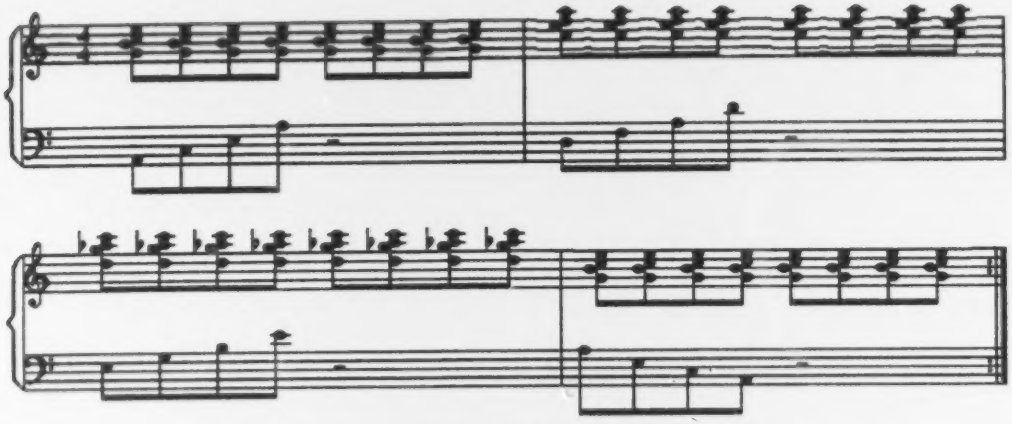


Fig 6b -version of "Abracadabra" in ninths

Working down the remaining chunks, the three chord alphabet is one of a limited number of possible restricted alphabets; an unrestricted alphabet would yield;

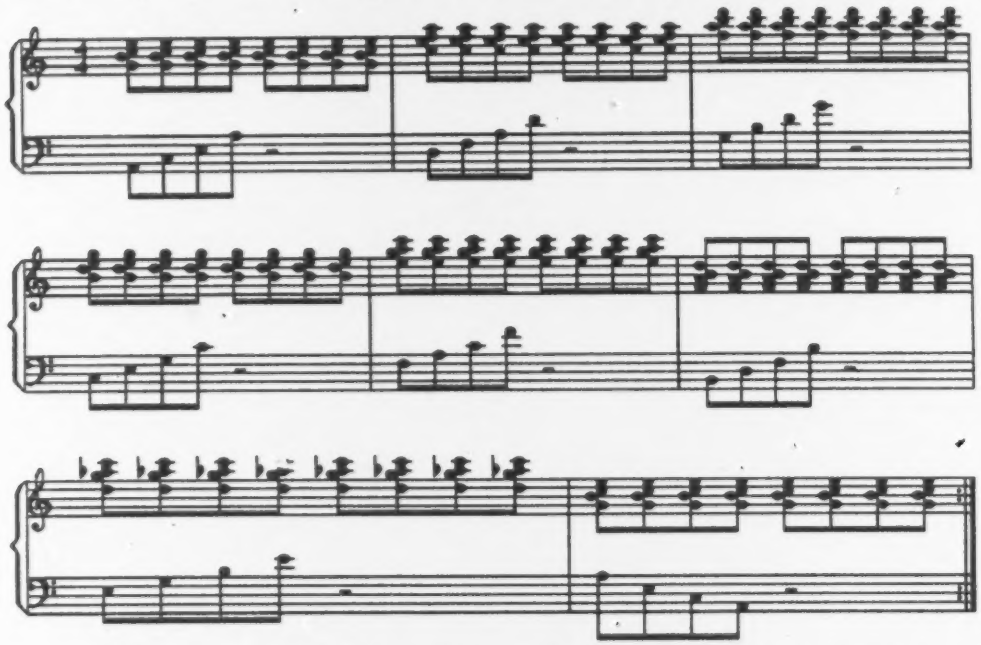


Fig 6c - A version of "Abracadabra" with chord alphabet restriction lifted

Finally, more variations can be obtained by replacing the chord sequence direction with another harmony space direction: for example, choosing an upward scalic trajectory, we would obtain;

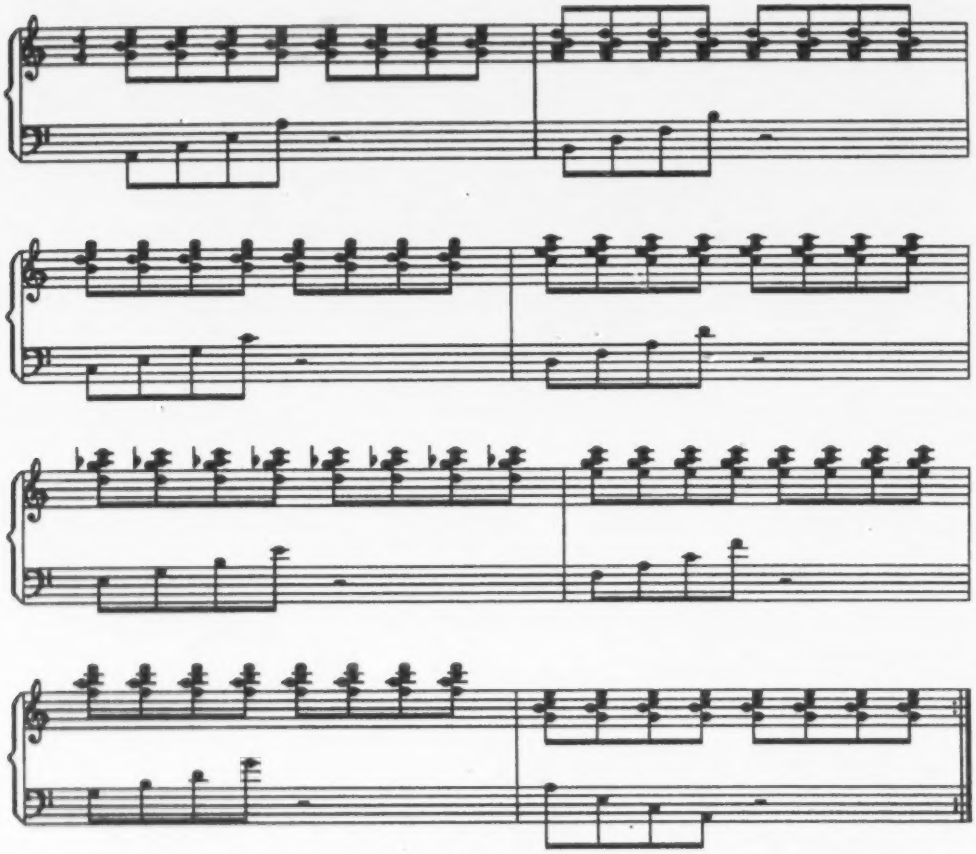


Fig 6d - version of "Abracadabra" with chord vocabulary restriction lifted and trajectory direction changed

Further variations could be made by requiring the bass line to play an arpeggio on a chord of a different arity, in the opposite direction and at a different metrical rate. On the other hand, we can consider the alternatives according to wider class membership; for example, the chord sequence direction chunk could be replaced with a procedure that makes no particular reference to the concept of Harmony Space direction and generates an arbitrary chord sequence; the motor rhythm could be replaced by any arbitrary rhythmic figure; the variant chord quality scheme could be replaced by any arbitrary variant quality scheme, the bass line could be replaced by an arbitrary bass line, and so forth. Such choices could in principle be indicated to a student by making inferences about class membership from a semantic net. The scenario of varying the chunks could be carried out using the same underlying formal mechanisms as the scenario of adding and subtracting chunks.

We are not claiming that any of the variations we have outlined have been achieved in particularly startling ways - although viewing a chord sequence as a emerging from the interaction of a harmonic trajectory and an alphabet restriction is a new approach. Neither do we claim that any of the outcomes have any particular musical merit (though the variation in fig 6c works well enough). However, several of the transformations, particularly the last two, seem to be highly suggestive of further possibilities. We will explore some of these possibilities in more detail in the section on musical plans.

This nearly completes our outline of the description of pieces at song level in the tutoring framework. In a full implementation of MC, each description of an aspect of fragment of a song or piece as a sequence (or tree) of chunks could be assembled into a package or "frame", with pointers or links from particular chunks and from the frame as a whole to related information at various levels (Fig 2). Some links could be to canned text, much as in hypertext links (Nelson, 1974), whereas others would be to nodes in a semantic net. This is not implemented, but is discussed in the further work section.

The reader may be wondering whether manipulation of the piece descriptions, as implemented in Harmony Logo, is simply Harmony Logo programming. This is essentially the case, but manipulating descriptions of pieces could differ from Harmony Logo programming in at least five ways. Firstly, the chunks need not be encoded Harmony Logo at all; secondly, the architecture need not require a novice to learn Harmony Logo programming as such, since the knowledge-based tutoring framework could constitute a front-end for manipulating descriptions of pieces; thirdly, the architecture need not require a novice to formulate descriptions from scratch, since it could centre around manipulation of existing descriptions of pieces; fourthly, the architecture gives a framework in which a knowledge-based tutor could support teaching strategies informed by user models (not implemented); and fifthly, piece manipulation could be combined with with style manipulation and constraint-based planning (implemented in prototype).

We now turn from the representation of pieces to the representation of styles.

3.3 Representation of styles

The next requirement for the tutoring architecture is to have knowledge available about a range of styles of music. We will use the words "genre", "style" and "dialect" synonymously and loosely to refer to partial descriptions in common to a range of pieces. For example, we might want descriptions of styles such as jazz, reggae, rock, blues, heavy-metal etc. There could be several descriptions associated with each style. We might even want detailed descriptions of "styles" relating to the prototypical patterns of particular performers, e.g. Michael Jackson, Donald Fagen, Stevie Wonder, Phil Collins etc. Note that for the purposes of this chapter, we are limiting ourselves to style as it applies to chord sequences and bass lines. Knowledge about pre-existing genres is important to composers in several ways: if not to conform to, then react against. The choice of style can have a large effect when choosing materials or strategies to employ. Stylistic considerations can also have a major effect on a novice's motivation. Irrespective of the choice of knowledge representation, crude but recognisable approximations of styles can be achieved using very simple means. For example, if any chord sequence is played in the following way;

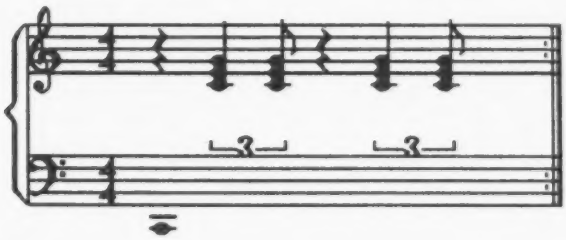


Fig 7 prototypical reggae accompaniment texture

it tends to sound as though it is in a reggae style. Clearly, this does not characterise the genre very deeply, but on the other hand, it is virtually universally applicable to any chord sequence, and so is useful as a caricature description. Similarly, a chord sequence played in sevenths in the following rhythm with the following bass pattern relative to the chord;

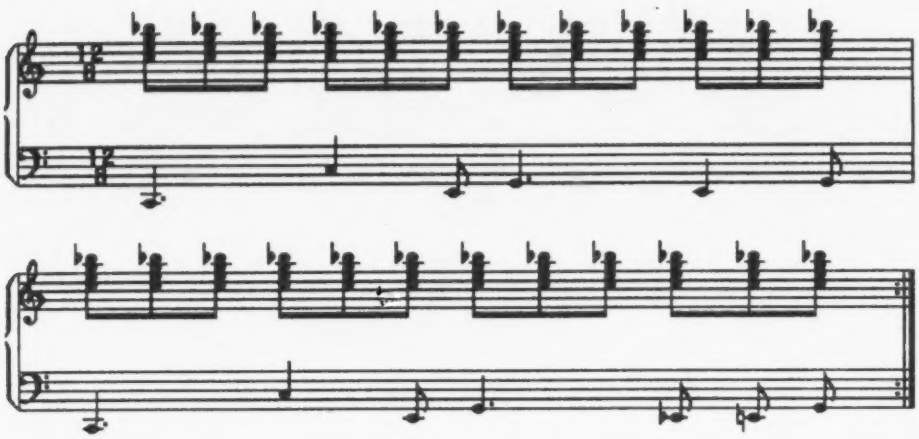


Fig 8: Prototypical Fats Domino arrangement (after Gutcheon, 1978, page 10)

sounds reminiscent of Fats Domino (Gutcheon 1978 page 10). This gets boring very quickly, but alternative patterns could be stored for variety. Even simpler means can caricature some styles: simply playing a chord sequence in ninths or sevenths in a suitable voicing scheme tends to sound "jazzy", especially if played with syncopation. Levitt (1985) demonstrates that it is possible to devise more detailed, typical characterisations of a genre but often with the trade-off that they will only apply to a narrower range of raw material: for example, a more detailed characterisation of a reggae style might only apply to pieces with a particular metrical structure, chord arity, harmonic rhythm and so on. As regards the representation of styles in the MC architecture, we have three key points to establish, most of which have already been established for us. We have to show that;

- Musical styles can be described in a variety of computational representations in forms that can be applied to stylistically alter existing melodies (or bass lines) and chord sequences.

- Constraints are a particularly good vehicle for describing styles.

- Style descriptions in more or less any formalism can be arranged in such a way as to be compatible with, and applicable to, the piece level and plan level descriptions of the MC architecture.

The first two points can be established merely by reference to the literature. Firstly, Levitt's (1985) work, together with notable work by Fry (1980), Jones (1984), Greenhough (1986), Desain and Honing(1986) and others has clearly established that important aspects of musical styles styles can be described in a variety of computational formalisms in a form that can be applied to transform existing melodies and chord sequences. Secondly, Levitt's (1985) work also established that constraint satisfaction is, in many ways, a formalism particularly suitable for describing musical styles. As Levitt (1985) demonstrates, such a formalism can characterise in a fairly natural way 'styles' which depend on interrelating chord sequences, melodies and rhythms.

This leaves us with only one point to be established; that style descriptions in more or less *any* formalism can be arranged in such a way as to be compatible with, and applicable to, descriptions at the piece level and plan levels in the architecture proposed in this chapter. This can be done in a very simple way by agreeing to use what Levitt calls 'pitch schedules' (essentially lists of notes and lists of note onset times) as an 'interlingua' between descriptions.[5] There is no problem about this from the point of view of the style descriptions: Levitt's programs take melodies (or melodies plus chords) as input described in just such terms. Programs by Fry (1984), Jones (1984), Greenhough (1986), Desain and Honing(1986), etcetera could easily be arranged to do the same. Neither is there any problem from the point of view of piece descriptions: note lists are the descriptions that Harmony Logo ultimately produces. Similarly, at the plan level (to be introduced later in the chapter) the constraint-based descriptions produce their output as Harmony Logo programs and hence ultimately as note lists. Thus, provided we borrow, from any of the above mentioned sources, or establish style descriptions in a form which can take note lists or chord lists as input, they will be compatible with, and applicable to, piece level and plan level descriptions in MC. The MC architecture is compatible with stylistic descriptions from any one of a number of sources.

In one sense, it is a pity that the constraint-based descriptions for musical plans (that

[5] In the case of chord sequences, the interlingua consists of pitch schedules (describing roots of chords) and lists describing chord qualities. The chord qualities may be expressed as semitone lists or in terms of symbolic primitives (the two can be made interchangeable by use of a look-up table).

we are about to introduce) have to be boiled down to note lists before they can be stylistically manipulated. It would be theoretically appealing if constraint-based descriptions of styles, chunks and plans could be freely intermingled as constraints. This is left for future research. In practice, pitch schedules make an adequate interlingua, since style transformations of the type we are considering are essentially a matter of decoration and elaboration, rather than deep transformation. For example, typical ways of altering styles include;

- changing texture (e.g. chord arity or quality)
- adding an accompaniment pattern
- imposing distinctive rhythmic patterns
- changing note or chord alphabet (e.g a change of mode or scale)
- adding decorative notes to an existing melody

Such changes can be carried out very adequately at pitch-schedule level, as Levitt's (1985) work demonstrates. To a certain extent, the MC architecture could allow intermixing of styles and chunks of pieces, both described directly as Harmony Logo code, if desired. Harmony Logo can easily characterise 'minimal' or caricature descriptions of styles, along the lines of those outlined above, by such means as changes of chord arity and rhythmic pattern. However, such characterisations do not run very deep. Informal experiments have shown Harmony Logo can be clumsy at describing styles in detail, since its predispositions for pitches to move in trajectories and events to occur at regular intervals need to be overriden rather frequently in some stylistic descriptions. This kind of informal experiment suggests that Harmony Logo is, as currently designed, a specialised language for harmony and education, rather than a general purpose music programming language. Still, this lack of suitability for describing styles deeply need not worry us, since we have shown that Harmony Logo descriptions are compatible with style descriptions from elsewhere, and since Harmony Logo pulls its weight in chunk and plan descriptions, and in the link with Harmony Space.

As in the case of descriptions of pieces, in a full implementation of MC, each description of an aspect of a style would need to be assembled into a package, or style "frame". Ideally, style frames should be organised in specificity/generality

hierarchies. For example, the reggae 'characterisation' given above could be extended into several successively more specific versions, which could be arranged in a tree.

In general, descriptions of pieces could be made both more economical and more informative by recording only skeletal information about the piece. The piece description would specify that, when the skeletal description had been re-constituted as a pitch schedule, parts of the the result should then be processed through specified style frames. In keeping with the MC approach, it would be acceptable if this produced an inexact reproduction of the original (though differences should be noted, and perhaps an audio recording of the original kept available on computer for comparison). Some pieces might need processing in their entirety through a single style frame; others might be patch-works of stylistic references. Even when encoding something as simple as the chord sequence of "Abracadabra", it might be illuminating to encode the momentary departure from rock norms (to the jazz ninth on the V chord) by recording that this single chord should be processed by a specified jazz frame, rather than literally specifying its constituent notes in the song frame. Apart from the advantage of representational economy, such a practice could serve the important educational purpose of recording links between pieces of music and styles that employ related musical resources. For example, a novice with little knowledge of classical music exploring "Abracadabra" might suddenly find the work of Debussy and Ravel of intense personal interest on learning via chains of such links (perhaps augmented by hypertext or semantic net links) that Debussy and Ravel did much to pioneer the voicings and compositional use of such ninths (Mehegan, 1985). For this and other reasons, it would be useful to record on each style frame back-pointers to the songs that make reference to a particular stylistic description. This could allow a student taking a *style* as a starting point to explore songs that were partially encoded in terms of that style. Similarly, it would be useful to record on song frames cases where a song was reminiscent of a style, even if it did not employ the relevant style directly. As in the case of descriptions of pieces, the collection of several alternative characterisations of a given style would be encouraged. Organisation of style frames along these lines is left for future work.

This completes our outline of the description of styles in the tutoring framework.

3.3.1 Applying the Paul Simon method to song and style levels

It is useful at this point to consider how the resources we have outlined so far - the chunked piece level descriptions and the style descriptions - could be used by a beginner limited to the Paul Simon method to compose new and interesting chord sequences. Given a collection of song frames of the sort we have illustrated, where descriptions of songs are already chunked in a way that accords reasonably well with musicians intuitions, a novice could select liked songs and then apply the Paul Simon method in a primitive way using subtractions, additions and substitutions of chunks. Any intermediate versions could be further influenced using a style description. 'Subtraction' simply corresponds to the process of creating new versions of the song with different combinations of chunks removed to find out which are vital, which are incidental and how they all work together. 'Addition' refers to the addition of chunks from other liked songs - although the results will not always be tenable. 'Substitution' refers to the substitution of a chunk with one from a different song. We saw examples of each of these operations in the earlier "Abracadabra" example.

3.3.2 Limitations of the Paul Simon Method

We will begin with a limitation affected by the technical means used to represent the knowledge, and then look at more general problems. This will help to motivate the next layer of the architecture.

In a version of the knowledge base that represented the piece chunks using low-level musical constraints, (as opposed to a version based on Harmony Logo) a process of combination of chunks could be very flexible but could be expected to produce useful results only some of the time. When adding chunks, a song could easily end up overconstrained and unsatisfiable. It could require the searching of a large space to satisfy a set of constraints. A further limit on the usefulness of the method would be that when making substitutions of chunks between songs, there would be no way in general of establishing an equivalence between chunks between two disparate songs.

In a version that used Harmony Logo as the base representation for song chunks,

fragments of program would be manipulated rather than musical constraints. However, Harmony Logo programs can set up a context for later Harmony Logo programs by manipulating, mode, rhythmic figures, chord vocabulary, trajectory direction and so forth. In some cases, of course, a Harmony Logo chunk might simply nullify the effect of a previous chunk, or produce a grotesque result.

In either case, when working at chunked song level, if we chose not to stick to obvious alternative values for chunks, we would run the risk of blindly swapping parts between perhaps quite dissimilar songs like an interspecies transplant surgeon. There would be no guarantee that any particular substitution would work. The lack of total reliability of the method need not constitute a major problem in this particular domain. Given the nature of the task, a process that succeeded even a small fraction of the time could be useful.

However, whichever way the song chunks are represented, applying the Paul Simon or intelligent transformation method at song chunk level is clearly rather limited. The starting point is always a particular existing song. The final piece will tend to be derivative from this song. The process of mutation is unguided, except by intuition: but a novice's intuition may be too inarticulate to do anything except evaluate steps once they have been made. A large search space can be explored more effectively if promising paths can be forseen in advance.

Without the right kind of knowledge being available to co-ordinate the vast number of possible choices, it can be easy to get 'stuck' in a space in which all 'nearby' variations are unsatisfactory. One of the roles of the plan level is to provide explicit knowledge about how the different musical materials interact, and to provide sensible musical goals to help the novice co-ordinate choices.

3.4 Musical plans

When chord sequences are analysed in Harmony Space, some superficially quite dissimilar chord sequences appear to conform to common underlying plan-like behaviours or schemata. Several such 'musical plans' have been developed in detail, but we will limit ourselves to discussing two contrasted plans informally; the "return home conspicuously" plan and the "moving goalpost" plan. We will begin with "return home", as follows.

Many chord sequences seem to begin by establishing a 'home area' that is perceptually distinguishable to a listener, after which they 'jump' elsewhere, and then 'return home' in such as way that it is evident to the listener that they are 'returning home'. We will refer to the postulated underlying plan as the 'return home conspicuously' plan or just the 'return home' plan. It is not being argued that composers necessarily plan out chord sequences consciously or unconsciously in this way. However, intuitively speaking, many musicians seem to agree that the general pattern described above is very widespread in music; it can occur with melodies, chord sequences, and to a lesser extent patterns of modulation. Clearly, there are several ways in which 'home' locations can be perceptually communicated in music. For example, in many ethnic musics, a melodic home area is communicated by means of a constantly sounded drone. In melodic traditions that use the diatonic scale, the "built in" major and minor tonal centres can act as perceptually discernable "homes". In modal harmony, the 'home area' may be marked by a pedal note, while in tonal harmony, the "built-in" major and minor tonal centres can act as 'home'. The degree of psychological validity of such plans is a matter for empirical research, but they are at least psychologically plausible. Generally speaking, musicians seem to agree that such 'musical plans' are a useful way of summarising information about chord sequences at a 'strategic' level. Here are several commonplace chord sequences that can be seen as instantiations of the 'return home plan';

"I got rhythm ' (Gershwin, 1930)	(Major) I VI II V I
"Johnny B. Goode" (Berry, 1958)	(Major) I IV V I

"Abracadabra"(Steve Miller, 1984)	(Minor) I IV V I
"The Lady is a tramp" (Rogers and Hart)	(Major) I IV VII III VI II V I (chorus)
"Street Life" (Randy Crawford, 1978)	(Minor) I IV VII III VI II Vdom7 I (chorus)
"Easy Lover" (Phil Collins, 1984)	(Aeolian) VI VII I
"Isn't she lovely" (Stevie Wonder, 1976)	(Major) VI II V I
"Out-Bloody-Rageous" (Ratledge, 1970)	(Dorian + modal pedal) III II I (coda)

For these pieces to qualify as cases of 'return home', each needs to be able to communicate a sense of 'home' location and a sense of movement 'towards home' to a listener. This is achieved in the various cases as follows. In each case, the piece communicates to the listener a sense of different 'locations' simply by employing different chord degrees. One chord degree is distinguishable by the listener as a 'home' location; in some cases this is communicated by tonal means and in other cases by modal means (e.g. using a modal pedal). In the tonal cases, movement 'towards home' is expressed in a perceptually apparent way as movement down the fifths axis towards I, whereas in a modal context, movement 'towards home' is perhaps most apparently expressed as scalic movement towards I. This is reflected in the examples by the fact that the sequences with a major/minor home area 'go home' down the fifths axis, whereas the sequences with a modal home area 'go home' scalically.

Given these typical resources, the 'return home' plan typically involves four steps: first of all the home area is established, secondly the piece takes up a location away from the home area, thirdly the piece moves homeward in such a way that it is perceptually obvious that the location is getting closer to home, and finally, home is reached and the piece finishes. The home area may or may not need to be explicitly stated at the beginning of the chord sequence, depending on what other musical resources are available to perceptually communicate its presence. [6] (In effect, this is a very simple example of sonata form (Jones, 1989).)

[6] In each of the last three examples, the home area is not initially explicitly stated. However, in each case we can think of the home area as being implicitly established in some other way. For example, in the Dorian example, the home area is communicated throughout by a pedal note, which provides a constant statement of the home area whatever the 'location' of the chord. Repetition, rhythmic and melodic means can also be used to establish a home area. Provided the home area is established in some way, these examples qualify as cases of the return home plan as defined above.

This basic skeleton plan can be instantiated with a great deal of variety depending on how the various elements of the plan are instantiated, and depending what external constraints are imposed, such as limited chord vocabulary, chord arity and so forth.

We can arrange a spectrum of cases of the "return home" plan corresponding to various choices in a tree, as in Fig 9. This is a simplified tree showing only a few of the plan variables, and only a few derived songs. The tree layout is perhaps slightly misleading in one sense, since amongst the group of elements in the plans to be instantiated, and amongst the group of external constraints, there may be no particular reason for any element to appear higher in the tree than any other element. The tree layout is a notational convenience for what is really a sparse n-dimensional array.[7] Since an n-dimensional array would be inconvenient to display graphically, the tree notation is a convenient way to help us visualise the space of chord sequences satisfying the plan. In the diagram, the names of plan variables are given in **bold** with an initial capital letter; choices of values for plan variables are given in plain text starting with lower case; and names of chord sequences are given in italics.

Musical plans such as "return home" have several potentially powerful educational uses. One such interesting use is the 'replanning' of existing songs. A prerequisite for this kind of use would be for each plan to have pointers to existing songs that

[7] The return home plan has variable elements such as ;

hometype	which may have value	tonal or modal
mode	which may have value	major, minor, dorian, ... etc.
initial non-home location	which may have value	any degree of the scale except I
route taken home	which may have value	any suitable direction

and so on. We refer to these variable elements in the plans as *plan variables*. The plan may also be considered to be affected by optional external constraints such as chord arity, restricted chord vocabulary and so forth. We refer to these as *external plan variables*. If the number of plan variables plus the number of external plan variables we wish to consider is n, then the space of chord sequences that satisfy the plan may be viewed a subset of the resulting n-dimensional array. Each dimension of the array is a plan variable or external plan variable. Not all points in the space will satisfy the plan, due to constraints on which combinations of values for plan variables are legal. In summary, the tree notation is a graphic convenience.

instantiate particular cases of the plan (i.e. pointers from plan to song level). Similarly, each song would need to keep pointers to each plan which it instantiates, and to record the corresponding values of the plan variables (i.e. pointers from song to plan level). Given these mechanisms, the student could start from a song of interest, and ask a series of "What if?" questions about the piece. The pointer system means in effect that each piece holds complementary or competing descriptions of itself or its parts as the unfolding of musical plans in particular ways. Having selected one such description for her chosen song, the student might pose, and get sensible answers to, a series of "What if?" questions as follows. For example, the student might ask how the chord sequence of 'I Got Rhythm', viewed as a case of 'return home', might appear if it had used a different mode, a different chord vocabulary or if it had employed modal materials instead of tonal materials. This can be viewed as 'navigating the plan tree' of fig 9. In some cases, this simply relates the original chord sequence to an existing chord sequence already present in the tree. (For example, we can view the chorus of "The lady is a tramp" as a replanned version of "I got rhythm" but with a different trajectory length for the journey home). However, in other cases, replanning a song in some specified way will produce a new chord sequence not corresponding to some known existing song.

But how exactly is a new chord sequence corresponding to a new leaf of the tree to be worked out in detail? For example, how is "I got rhythm" to be replanned using modal materials? In general this requires some musical knowledge about how to use particular materials to achieve particular effects: for example, to replan a song using a modal as opposed to a tonal hometype may require knowledge about how to make modal hometypes perceptually apparent, and knowledge about the kinds of harmonic 'route' home that tend to emphasise perceptually the existence of a modal home area.

An expert musician might tend to know this kind of thing anyway, and could perhaps use verbal descriptions of the plans and plan trees as "possibility" indicators, to give ideas for new interesting chord sequences that she might otherwise tend to overlook. However, this kind of knowledge is precisely one of the things that a novice might be expected to lack. The implemented musical planner makes use of formally stated versions of the plans that record just this kind of knowledge.

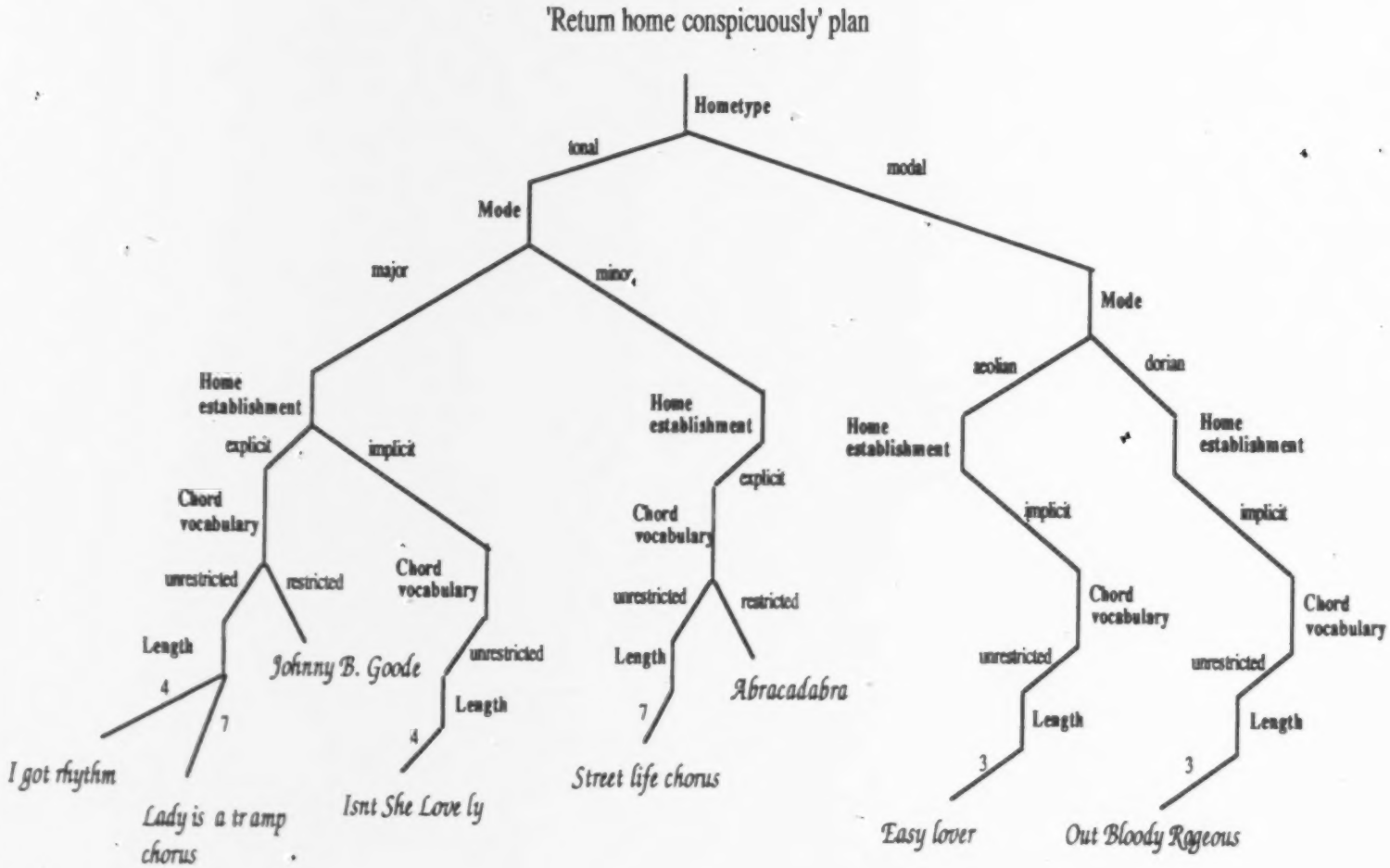


Fig 9 Tree of songs instantiating the "return home conspicuously" plan

We will give some informal examples of the kind of knowledge we have been talking about in our next example, based on the "moving goal post" plan. We have seen how the "return home" plan is good at relating diverse commonplace chord sequences, but it gives little impression of the depth of transformation that the musical planner is capable of performing (while keeping the result musically intelligible, interesting and related to the original). The next example, using the "moving goal post" plan shows a deeper transformation based on an underlying plan, as well as casting light on the notion of a plan based on a single existing song. This example is as follows;

3.4.1 Plans based on a single chord sequence

The notion of an underlying musical plan may make sense even when the plan is identified only in a single existing song. For example, consider the analysis of the chord sequence of "All the things you are" (fig 10), as discussed in chapter 5. The analysis was framed in terms of the movement of a 'goal post' that synchronised metrical and tonal arrival. This analysis is already couched to a large degree in high-level, 'strategic' terms. It is tempting to try to generalise this description and cast it abstractly as a formal plan, and then find out whether we can use the plan to 'recompose' the song in various different forms while retaining something of the effect of the original chord sequence.

Amin7 Dmin7 G7 Cmaj7 Fmaj7 B7 Emaj7 Ema6

Fig 10 Original chord sequence - first 8 bars of "All the things you are", harmonic tempo one chord per bar. (Kern Hammerstein)

To this end, we might hypothesise a 'moving goal post' plan on the basis of the former analysis, which we can outline informally as follows;

- In order to satisfy the plan, a chord sequence should communicate the

presence of a definite, perceptually distinguishable home area. This could be done in a variety of ways, for example by means of a pedal note in the modal case, or in the tonal case, simply by the use of triads in typical tonal progressions with no accidentals and no shifting of key .

• The sequence should begin with a perceptually obvious, regular movement directly towards the home area, for example, a trajectory down the fifths axis in a tonal case.

• The plan assumes the expectation on the part of the listener that obvious, regular movement towards a home area will normally be synchronised with arrival at a strong metrical boundary. (This would tend to be expected as the norm in many musical contexts, for example in jazz chord sequences).

• The chord sequence should communicate the location of strong metrical boundaries at appropriate points in the sequence, for example by the use of strong and weak emphases on chords to suggest the presence of a strong metrical boundary at, say, the 4,6, 8 or 12 bar mark.

• The harmonic progression should continue *past* the home area with the original regular movement. The metre and regular chord sequence should be such that the harmonic progression will not reach a home area at the next strong metrical boundary after this overshoot in the absence of some intervention.

• After the tonal goal has been initially passed, there should be a modulation such that the continuation of the original regular harmonic movement produces a synchronised arrival at the new "home area" and at a strong metrical boundary.

3.4.2 Replanning an isolated chord sequence

How exactly could a formal version of such a plan be used to "replan" the chord sequence of "All the things you are" in a musically intelligent way? For example,

what might a modal recomposition of the original chord sequence look like? These questions have been investigated in detail using the implemented musical planner and a formalised version of the above plan. We will content ourselves for the purposes of this chapter by outlining how a modal transformation might proceed, in informal terms, for the first eight bars.

Firstly, the plan requires a home area that is perceptually distinguishable. Substituting a modal home area for the tonal home area of the original requires a piece of musical knowledge for effectiveness. Namely, that to make a modal home area perceptually obvious to a listener, explicit measures (such as a modal pedal) must be employed, in contrast to the case with the 'natural' major/minor home areas. Hence, the transformation might begin with the picking of a typical mode such as the Dorian mode, and the specifying of a modal pedal note in the bass.

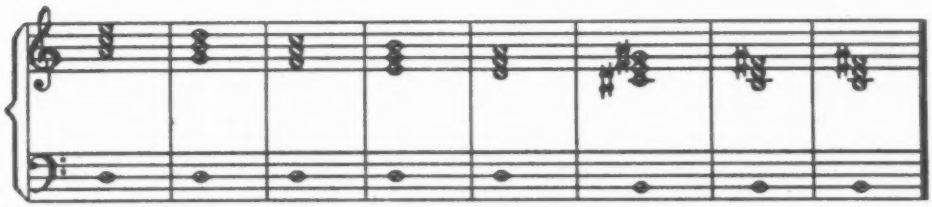
The next requirement of the plan is for an obvious, regular movement towards the home area. In the original piece, the movement proceeds down the fifths axis, which gives a very strong sense of regular movement towards the tonal goal. To transform this most naturally into a modal version requires another small piece of musical knowledge. Namely, in the case of a modal home area, a sense of strong movement towards the home area is perhaps better expressed by *scalic* movement towards the home area.

The next and final requirements of the plan concern synchronisation of arrival at tonal and metrical goals. To keep matters brief in this informal account, we will simply 'borrow' the relevant timings from the original in a rather mechanical way. This way of tackling the synchronisation part of the task is not the way a composer would be likely to work. However, it will serve our purposes as an expositional device for reasons of brevity and definiteness. In the original piece, the timings are as follows: the strongest metrical boundary is at the eighth bar, the piece reaches 'home' on the fourth event, the modulation occurs between the fifth and sixth events[8], and the sequence arrives at the new home point on the seventh event. The eighth event is used to reiterate the home event and emphasise the arrival at the home area, making use of a common device in tonal music.

[8] In the first eight bars of "All the things you are" there is effectively one harmonic event per bar.

To put these timings to use to complete the replanning, we take the further small piece of musical knowledge that Dorian sequences tend to work better if they move *down* the scale towards the home area (in contrast with, say Aeolian sequences that tend to work better if they move *up* the scale towards the home area).

Putting everything together, we plan the progression by working back up the scale, to find out where in the scale a downwards descending Dorian progression would have to start to hit the home area on the fourth event. The answer is G (in D Dorian). If we now mimic exactly the timing of the original piece: "..the strong metrical boundary is at the 8th bar, the piece reaches 'home' on the fourth event, the modulation occurs between the fifth and sixth events, and the sequence arrives at the new home point on the seventh event.."; we arrive at the following chord sequence (fig 11) ;



(D dorian) Amin G maj Fmaj Emin Dmin

(B Dorian) C#min Bmin Bmin

(with a D Dorian pedal in the bass, changing to an B dorian pedal on the C# minor)

Fig. 11 Example modal replanning of "All the things you are"

Note that the modulation has been interpreted in modal terms in the modal context. The reader may judge whether or not this has a similar effect to the original sequence. In fact this is not a particularly elegant modal replanning of this chord sequence. This is only to be expected, since for the purposes of brevity of exposition we used a mechanical, inflexible way of planning the timing of the chord sequence. As we have already noted, a composer might have carried out this part of the task in a far more intuitive fashion, and the implemented planner can carry out this part of the task much more flexibly.

Still, we have at least demonstrated informally the reasoning behind one possible example of a modal recomposition of "All the things you are", guided by relevant musical knowledge.

3.4.3 Further work on musical plans: testing musical plans by experiment

As in the "return home" examples, such plans might be used principally as ways of exploring how songs achieved their effect, and as ways of carrying out the Paul Simon method at plan level. However, an interesting alternative perspective on the framing of musical plans is as follows: Let us suppose that we replanned several versions of a song, and these turned out to be different on the surface from the original, but were judged intuitively to produce a similar effect. We would then tend to feel that the original analysis and its formalisation had been justified to some degree by the experiment. To a limited extent, we would have an experimental tool for testing our musical analyses. Conversely, if replanned versions failed to capture the same effect, this might give us concrete evidence of the flaws in our original analysis, in its formalisation, or in our conception of how musical materials interacted. This could either lead to the refinement of the original analysis, or its rebuttal.

3.4.4 Representation of musical plans

Most but not all of the musical plans that we will be presenting are easy to visualise when presented in terms of Harmony Space concepts: indeed, this is how many of them were originally identified. For this reason, the musical plans are represented in terms of constraints applied to relationships expressed in harmony space primitives.

3.4.5 Summary of educational uses of the plan level

In outline, musical plans have three broad classes of educational use: as a way of conceptualising existing chord sequences, as a way of relating them in families and as a way of composing new sequences. In particular, musical plans may be used to pursue the Paul Simon Method at plan level. An unexpected advantage is that since the return home plan can be framed intuitively in terms of everyday, extra-musical notions such as location, 'home' and spatial movement, it can be communicated, at least in outline, to novices with no specialised musical skills at all but with access to

a Harmony Space interface.

When learning to compose, the temptation to imitate existing works, or conversely the fear of imitating existing works can hinder a beginner's attempts: having clear strategic goals to pursue within chosen constraints can be very helpful. At the same time, the plans are not intended as a set of normative patterns to be learned for their own sakes, but as a way of organising and illustrating knowledge about how musical means can be made to serve musical ends: it is hoped that like Wittgenstein's ladder (1961/1921), the plans would be discarded once the novice had learned how to use musical materials to satisfy workable musical goals.

4 Intelligent Tutoring Systems' perspectives on the framework

To give some perspective on our framework for a tutor, as it has been so far presented, we will quickly review some of the key points of Self's (1988) summary of recent developments in ITS architecture and describe how our framework relates in several ways to these wider trends .

- **The need for multiple representations of expertise.**
  Research on Gu: lon/Neomycin (Clancy, 1987) demonstrated clearly that competent explanation in different contexts may require more than one model of the expertise in question. One single model may be able to perform domain tasks adequately, but is unlikely to serve as a basis for adequate explanation in all contexts. A key feature of the conception of MC is the use of a series of multiple representations. Firstly, the student may hear the notes of a piece aurally, or see its trace visually as an animation in harmony space. Secondly, each song is also represented in a chunked format, where each chunk has some intuitive meaning demonstrable in a microworld: a given song may be chunked in more than one way to reflect different possible analyses of the piece. Thirdly, pieces may be represented as some combination of a simple outline and various styles. Finally, a piece may be represented as an instantiation of a musical plan or schema: some pieces may have several alternative descriptions in terms of different plan schemas.

- **A focus on strategic and metacognitive skills.** A number of recent

ITS systems (for example, Algebraland (Foss, 1987) and EPIC (Twidale, 1989)) require or help the student to specify some activity in goal oriented or strategic terms. This can benefit in two directions; aiding the learner to develop appropriate metacognitive skills and helping the ITS to build up an appropriate student model. A key feature of the MC architecture is that it allows a student to work in goal-oriented or strategic terms. If the student starts from an existing piece, the architecture encourages the student to look at the piece as being goal oriented according to several competing plan descriptions. Alternatively, the student can take a schema or musical strategy as a starting point. Note, however, that the recognition of plans in songs that the student drafts from scratch has not been explored, nor the provision of means for the student to modify or create new musical plans. There is a strong emphasis in the architecture on metacognitive skills. Rather than concentrating on the low level elements of chord sequences, the student is encouraged to consider the ends that chord sequences serve and the kind of high level considerations that can influence their shape.

• **New interaction styles** While early ITS research used a variety of interaction styles, e.g. Socratic dialogue, case-study based teaching, mixed initiative dialogue, coaching, etcetera, these tended to have a common aim of "remediation". In other words, each style tended, however tactfully or sparingly, to be bent primarily on transferring to the student an item of domain knowledge. By contrast, a more recent system, DECIDER (Bloch and Farrell, 1988) used its domain knowledge to pick out evidence relevant to a student's hypothesis that might lead the student to elaborate or revise her opinions, but without reference to any preconceived target hypothesis on the part of the system. On a related note, Gilmore and Self (1988) experimented with machine learning techniques in a collaborative tutor aimed not at transferring pre-existing expert knowledge to the student, but at collaboratively discovering new knowledge. The framework of this chapter is certainly not remedial in the sense we explained earlier. Although the plans are, in one sense, normative, it need not matter whether the student 'learns' any of the plans or any of the songs. Successful outcomes are essentially of two types: firstly that the student composes interesting new

chord sequences - though the system has no targets for the exact form these may take, and secondly that the student comes to appreciate that chord sequences have a strategic dimension, and that various low level means may be used to serve various high level musical ends.

5 Partial implementation of MC

The following components of MC are implemented in prototype form;

- song level descriptions,
- plan level descriptions,
- the song interpreter,
- the plan interpreter,
- the Harmony Space microworld,
- the song description default providing mechanism.

Links between the different levels are left for further work. These components as they stand appear to form the basis of a richly interacting educational toolkit. A toolkit based on these prototypes appears suitable to be used with the guidance a human teacher, or embedded within a hypertext-like system with supporting nets of textual information. However, to build a knowledge-based guided discovery learning system for music composition with *active* teaching strategies initiated by the system would require these core components to be augmented by further components such as a semantic network of musical concepts; user and curriculum models; stereotypical user models; appropriate teaching strategies and adaptive microworld exercises. Under the heading of "further work", we will outline ways in which this work could be approached. The primary concern of this chapter has been to establish an architecture that makes a tutor feasible at all: the aim has been to devise ways of characterising strategic considerations about chord sequences and find characterisations for the song, style and plan levels that are suitable for supporting novices' needs.

6 Further work: supporting active teaching strategies

Let us consider a novice who is musically uneducated and wishes to compose music but doesn't know how to set about it, or someone who is musically educated but

can't compose. A human tutor might begin by focussing on three questions;

- What sort of music do you want to compose?
- What sort of music do you like or are you interested in?
- What musical skills and knowledge do you have?

Answers to the first question could be given given in terms of particular pieces, styles, composers, musical materials, musical techniques or some combination of these. This helps establish a provisional common goal or agenda for the teacher and novice, albeit in a crude way. The teacher can use this information to infer target musical techniques and materials that may need to be learned. Asking the student what she likes can help to identify a pool of material liked by the student that can be used for illustrative purposes while moving towards some chosen goal. For example, if the answer is given in terms of a particular composer or group, whenever some concept or technique needs to be taught, a corpus of work is available to search for examples of the concept or technique in a form likely to motivate and engage the learner. Finally, the question about skills and knowledge allows the teacher to identify strategic gaps in knowledge or technique and to generally assess what the student already knows.

On the basis of answers to these questions, renewed as appropriate, a human teacher could dynamically construct a teaching plan and exercises tailored to her pupil's likes and capabilities. Of course, the teacher might well be influenced by a notional pre-existing 'core curriculum' of generally useful compositional strategies, techniques and materials. Many broad strategies for teaching would be possible. One teaching strategy might be to provide intermediate goals based on the amateur's self - expressed preferences in music as given in answer to the initial questions. For example, a human teacher might think "The student wants to compose like (say) Steely Dan, so she is going to need to know about voicing 9th, 11th and 13th chords, about using cycle of fifths progressions, about transient modulation, about use of rhythmic motifs, melodic trajectories, and so forth. She doesn't know about any of this yet. We could start with a simple blues, and then look at some simple Jazz standards using cycles of fifths before getting on to Steely Dan". This approach appears reasonable, provided the student is willing to settle for the intermediate

steps as enjoyable substitute goals for the time being, or is convinced of the relevance of the intermediate steps to reaching her goal. A contrasting strategy might be to say " Composing like Steely Dan is beyond this student's capabilities as yet, but by playing with these simple components she could get a very rough approximation right now". A reasonable combined strategy might be to offer the amateur a continuing free choice between these two strategies.

A first approximation of this teaching strategy appears to be well within the range of existing Intelligent Tutoring Systems techniques using MC as a knowledge-base: although laborious to implement for this domain. The task could be approached as follows. The first task would be to construct a domain-specific semantic net of concepts required. This would be little different in principle from semantic nets for other subjects, such as geography (Carbonell, 1970b). Broadly speaking, the role of the semantic net would be to interrelate musical concepts used in the song, style and plan knowledge sources tc each other, and to relate these concepts to the concepts of standard music theory. In the present case, we would need to compile a network of concepts used in song, style and plan descriptions (including, for example, low-level concepts such as note, chord, chord-quality, chord-arity, chord sequence, key, metre, harmonic trajectory, mode, degree of scale, pitch, etc, as well as higher level concepts such particular songs, styles and musical plans). The relationships between the concepts would then be recorded using relations such as specialisation, generalisation, 'similar-to', and possibly domain specific relationships. Detailed examples of the use of semantic nets in ITS applications can be found in Carbonell (1970) and Wenger (1987), *passim*. See Baker (in press) for a semantic net used for tutoring in a musical context. Having constructed such a semantic net, which typically is very laborious, answers to the three "user-model" questions outlined above can be used to construct three distinct overlays representing the learners goals, preferences and skills/knowledge. A simple technique to make the creation of a complex user model of this sort viable is to set up a set of stereotypical user models (Holland, 1984), (Rich 1979), that allow a user to indicate her abilities or preferences in terms of a set of stereotypes. In addition to the three overlays already mentioned a core curriculum overlay could be used to indicate items that are particularly important or useful to learn.

Given a semantic net, a set of user models and a core curriculum, teaching strategies

could be used to find paths dynamically through the knowledge-base of songs, styles, chunks and plans. In essence, such paths would work towards an agreed goal via materials that the user liked, introducing not too many new concepts at once. At each step, the knowledge base could support a rich range of specific teaching actions such as

- invitations to carry out the Paul Simon Method with particular songs, style and plans with special attention to particular chunks
- presentation of explanatory text associated with particular concepts
- presentation of adaptive exercises associated with particular chunks or abstract plan steps to carry out in the microworld instantiated with values drawn from particular songs or plans. Such exercises could include analysis, attempted reproduction of an unseen piece, modification of an existing piece and so on.

This completes the section on further work on the MC framework.

7 Problems and limitations

The "knowledge" in the knowledge base is not intended to have the status of items of certified music theory or accredited music psychology. Rather, the chunking of songs, characterisations of styles and musical 'plans' are intended to have the heuristic value of hints and ideas that musicians pass on to each other, particularly in informal situations. Hence, there has been some simplifications of musical theories for educational reasons, and some musical technical terms have been used in unusual ways. In any case, psychological evidence about the process of composition is very sparse (Sloboda, 1985), and detailed psychological or music-theoretic discussions about strategic considerations for devising chord sequences in popular music is almost non-existent (for some of the nearest approximations, see Cannel and Marx (1976), Making Music (1986), Mehegan(1959), Cork(1988), Steedman(1984) and Citron (1985)).

Given suitable microworlds for rhythm and melody, and representation of melodic and rhythmic aspects of pieces at three levels, (perhaps represented in a constraint-

like formalism), an interesting topic for further work would be to try to integrate the treatment of melody and rhythm within the MC framework.

8 Summary and conclusions

We have presented a framework for a knowledge-based tutoring system MC to help novices explore Harmony Space. The framework is aimed at providing support for composing chord sequences and bass lines. In creative tasks there is no precise goal, but only some pre-existing constraints or criteria that must be met: in *open-ended* creative domains such as music, the creator may now and again drop any pre-existing constraint when she considers some other organisational principle to take precedence (Sloboda, 1985). The standard means of representing domain expertise (typically as rules) sits uneasily in a domain where the rules may be rapidly changing according to principles that may not be known in advance. An approach to thinking about creativity characterised clearly by Johnson Laird has been used to devise an ITS framework for dealing with open-ended domains. It appears that MC is flexible enough to support a very wide range of learning and teaching styles.

The framework allows multiple alternative views of pieces to be explored by students. Pieces are represented at three levels: at surface (note) level in a chunked form; with reference to common musical styles; and with reference to strategic plan-like or schema-like descriptions. At each level, alternative descriptions may be held. The three levels give multiple views of existing pieces, but they also allow aspects of songs to be stripped down, mixed, played in different styles or 'replanned' at different depths using new materials. Work on the plan level involved developing an original approach to describing strategic considerations about chord sequences, based on Harmony Space concepts and simple psychological considerations.

Although the plans are, in one sense, normative, it need not matter whether the student 'learns' any of the plans or any of the songs. Successful outcomes are essentially of two types: firstly that the student composes interesting new chord sequences - though the system has no targets for the exact form these may take, and secondly that the student comes to appreciate that chord sequences have a strategic dimension, and that various low level means may be used to serve various high level musical ends.

When songs are represented in the form of musically meaningful "chunks", studying the effect of mutations is more likely to be illuminating for novices than altering individual notes. Example chunks have been exhibited based on Harmony Logo primitives, which seem close to common intuitions about pieces and similar informal terms used by musicians.

Musical plans have been illustrated informally. A simple fall-back user-led strategy which the student could use to navigate through a series of transformations in the knowledge-base, the 'Paul Simon' method, has been presented.This method is one example of a non-prescriptive, idiom-independent general-purpose composition strategy that can make use of any knowledge sources. Teaching strategies to provide individually relevant guidance are not discussed.

Existing or new pieces derived from the three layers could be explored aurally and visually in the Harmony Space microworld. The case studies and the plans could be used as material to guide the exploration of Harmony Space. Preferences, knowledge and goals could be used to constrain search. Reciprocally, Harmony Space could be used to elucidate the concepts used in the piece chunks, style descriptions and plan descriptions. MC provides a framework within which a complementary explanatory process could be carried out in both directions using canned text and set piece exercises in combination with more flexible adaptive exercises. The MC framework constitutes a good basis for a Guided Discovery Learning system (Elsom-Cook, in press). It can also be viewed as something between a framework for an aural, generative hypertext system and a song synthesiser.

It is argued that in the light of Johnson-Laird's (1988) characterisation of creativity, it is particularly appropriate, where possible, to represent musical knowledge for compositional purposes in terms of constraints. On the other hand, many musical plans and harmonic concepts seem particularly easy to visualise in harmony space. Hence partly to aid ease of communication with novices, and partly for practical reasons, prototypes of the three representation levels are implemented in terms of Harmony Space primitives, but in the case of the strategic or plan level with a

constraint satisfaction system implemented on top of the primitives. The song level is implemented in prototype as a Harmony Logo interpreter, (which indeed was the main motivation for implementing the Harmony Logo interpreter).

Many features of the MC architecture may be applicable to other open-ended creative domains. In particular the guiding metaphor of the Johnson-Laird characterisation of creativity, not previously used in an ITS context, use of constraint satisfaction as a representation, The guided discovery learning approach, the notion of object level, style level and strategic level of description, the Paul Simon or "intelligent transformation" method and the outline ideas for active teaching strategies appear open to use in other creative domains. However, the particular form of the Harmony Space microworld, and the particular way its concepts permeate and give pedagogical leverage on the conceptual and strategic primitives is specific to the domain.

Part IV Conclusions

Chapter 9: Novices' experiences with Harmony Space

1 Introduction

In this chapter we report on the experiences of some musical beginners carrying out various educational activities using Harmony Space. The aim of the chapter is to illustrate and illuminate the use of Harmony Space on the basis of a qualitative investigation.

We report on five subjects, aged from 8 to 37, who had single sessions with the prototype version of Harmony Space. All of the subjects had negligible previous musical training and negligible musical ability. The sessions lasted for between 30 minutes and two and a half hours, one session being broken by a 45 minute meal break. The author acted as instructor. The sessions were video-taped and audio-taped. Examples of the following activities were carried out;

- playing the chord sequences of a variety of existing pieces,
- composing new chord sequences in a purposive manner,
- modifying existing chord sequences in musically 'intelligent' ways,
- playing accompaniment for a singer,
- harmonically analysing modulatory chord sequences,
- discussing elementary aspects of music theory.

1.1 Limitations of the prototype affecting the investigation

The prototype as implemented has many limitations that made it difficult to use for the investigation, which are documented in appendix 1.[1] These limitations reflect

[1] In particular, although the prototype exhibits all of the key features of the core abstract design, the prototype is non-DM in secondary control functions; it features modes, inadequate labelling, treacle-like cursor action and hidden states.

the original reason for its construction. The prototype was originally implemented for the purposes of demonstrating the coherence and feasibility of the Harmony Space core abstract design: it was not intended to carry out experiments with novices. The decision to carry out an illuminative study was taken at a much later stage. Despite the limitations, the prototype proved adequate to support a useful qualitative investigation.

2 Method

The study was carried out as follows. From a group of around thirty potential subjects, some seventeen candidates were initially selected as having minimal musical ability or training. After detailed questioning, twelve of the candidates were rejected due to an ability at some time to play an instrument, or due to childhood musical training (irrespective of whether this had been forgotten). Only five candidates, with effectively no knowledge of music theory, and effectively no ability to play musical instruments remained. These five subjects were allowed to go forward to the investigation. Details of the subjects' backgrounds are discussed individually below. The instructor took the five subjects one at a time and taught them by means of informal demonstrations, explanations, exercises, and theoretical discussions to carry out as many of the tasks described below as possible. To compensate for the lack of any labelling on the makeshift prototype, three diagrams showing the note circles appropriately labelled (figs 1, 2 and 3), were available for the subjects to refer to. (This introduced nothing not already an integral part of the parsimonious Harmony Space core abstract design: all three kinds of labelling should be available in any version of Harmony Space designed for practical use.) The selection of tasks that individual students were asked to perform varied according to the time available, the subjects' ability and interests, and the tasks remaining to be illustrated. The tasks fell into groups as follows;

- To accompany, playing the correct chords, sung performances of various songs, for example;

  Transvision Vamp's "Baby I don't care", The Archies' "Sugar Sugar", The Kingsmen's "Louie, Louie", Steve Miller's "Abracadabra", The Beatles' "Birthday", The "Banana Splits Show" theme song, Randy Crawford's "Street Life", Bob Dylan's "All along the Watchtower",

Michael Jackson's "Billie Jean", Phil Collins' "In the air tonight", George and Ira Gershwins' "I got rhythm", Stevie Wonder's "Isn't she lovely", Rogers and Hart's "The lady is a tramp" and whatever else was known to the student and seemed appropriate to illustrate a point,

• To harmonically analyse performances of the chord sequences of parts of pieces containing modulatory passages such as;

- Mozart's 'Ave Verum Corpus'
- The Rolling Stones' 'Honky Tonk Women'[2]

• To learn simple strategies for composing chord sequences such as

- return home
- cautious exploration[3]
- return home with a modulation (to be explained below)
- modal harmonic ostinato

• To perform 'tritone substitutions' (see chapter 5, section 4.4.3) on examples of a class of jazz chord sequences,

• To play on request and recognise a variety of abstractly described chord sequences, either diatonically or chromatically, for example:

- major 'three chord tricks' and minor 'three chord tricks' (chap. 5, 4.3)
- cycle of fifths progressions from major tonic to major tonic
- cycle of fifths progressions from minor tonic to minor tonic
- scalic progressions to and from modal centres

---

[2] 'Honky Tonk Women' does not contain a true modulation - it merely 'borrows the resources' of another key. But it turns out that both concepts are easy to explain and distinguish to novices using Harmony Space - though apparently difficult to explain by conventional means.

[3] 'Cautious exploration' is a musical plan in the same spirit as 'return home' and 'moving goal post'. The essense of the plan is that almost any chord sequence that starts on the tonic, moves to 'nearby chords' (in one of various naive Harmony Space senses) and then finishes with a 'sensible' cadence will make a workmanlike chord sequence. This plan, and several others, has been analysed carefully with examples from the literature and implemented as a constraint based computer program.

- To locate, recognise and distinguish examples of fundamental harmonic concepts, for example:
  - To locate and play the major and minor tonics
  - To recognise and control nearby modulations
  - To distinguish major and minor triads in various keys

To learn about tonal centres and chord construction

- To learn the rationale for the centrality of the major and minor tonics
- To learn the rule for scale tone chord construction (in any key)
- To learn the link between scale-tone chord quality and root degree

Further details about the tasks are given in the descriptions of the sessions. Not all tasks were introduced to all subjects. In some cases, where time was short, the tasks were presented as mechanical tasks, without much attempt to explain their musical significance. This was done solely for the purposes of the investigation, to make the best use of limited time. In other cases, every effort was made to make the student fully aware of the musical meaning and context of the task. No attempt was made to use standardised forms of words with each student; to the contrary, the demonstrator acted as an engaged participant. This was done for the following reasons;

- It is unclear how novices with negligible musical training could be taught to carry out tasks such as those listed above (in arbitrary keys) in two hours by any other method, irrespective of whatever form of words a helpful teacher might use.

- There is no claim that novices can learn to use Harmony Space effectively without initial guidance.

- The purpose of the investigation was illuminative.

We will now describe the five sessions in the order in which they were carried out.

3 Session 1: YB: Harmonic analysis by mechanical methods

YB is a 26 year old first year research student in information technology and education with a Computer Science background, from mainland China. YB has been in the West for less than two years. He has no formal knowledge of, and very little exposure to, Western music, once memorably asking "Who are the Beatles?". YB does not play any musical instruments and has had negligible musical training (even in China with respect to Chinese music). He had not had any previous exposure to the research reported here.

The session lasted about half an hour. This session was a preliminary trial, not originally intended as part of the investigation. Consequently, it was the only session not to be videoed or tape-recorded. Notes were made immediately after the session. The intention was to discover whether students could make sense of the display and thus perform a harmonic analysis task.

YB was given a five-minute informal practical and theoretical introduction to the interface, being shown the relationship between degree of scale, chord quality and modulation. There was no attempt to explain the analysis task fully before it was carried out - due to the preliminary nature of the session and lack of time it was treated as a mechanical task. This violates the spirit of Harmony Space, but was a practical necessity for the purposes of the investigation.

The task, in general terms, was as follows (we will give more precise detail as we go along). The investigator played a modulatory chord sequence on the interface. YB was asked to watch while it was played and identify the names of each chord and the direction of any key window movement. This was treated as a game in which the investigator made a 'move' and waited for the student to respond with a chord name or direction of window movement before making the next move. (The situation may seem a little unnatural for harmonic analysis, but then perhaps so is sitting down with a pencil, manuscript paper and score.)

As already indicated (since the prototype is unlabelled), YB had a diagram (fig 1) showing a labelled version of the key window pattern. To make sure that YB interpreted the screen movement from the screen and not the cursor keys, the cursor

keys were hidden from him. YB was played the first few bars of the chord sequence of the Rolling Stones "Honky Tonk Women" as follows, played in triads, in root inversion, in close position, moving the key window to the dominant just before the #3 chord. Each chord was held until YB had responded in some way.

C F C D#3 G

Unsurprisingly, YB identified each chord more or less correctly (we will note one mistake he made in a moment) with respect to its position in the key window, simply by referring to the labels on the diagram, and pointing to indicate the direction of modulation. YB's correct analysis was as follows:

I IV I (North Easterly movement of key window) V I

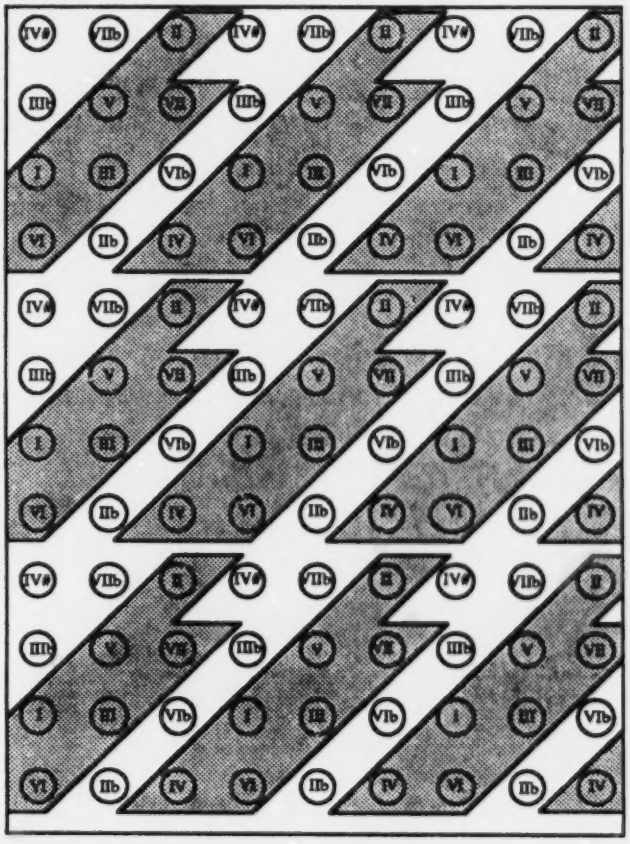


Fig 1 Harmony Space with chord function labelling

In the conclusion of this chapter, we will comment on this procedure in more detail

and illuminate the relationship that it bears to conventional harmonic analysis. In descriptions of later sessions we will show how some subjects successfully carried out more refined versions of the same task, with a reasonably clear understanding of what they were doing. We will also show how some of the other students were taught to use more traditional vocabulary for describing modulations. But first of all, we will cover two points of detail. Firstly, YB mis-identified a I chord as a III in his first attempt because, performing this as a mechanical task, he didn't realise that he had to look where the cursor was to name each chord - the cursor is not strongly visible on the prototype. (Due to time constraints, no effort was made to teach YB the rule for chord construction, which could have enabled him to label chords without needing to note the position of the cursor.)

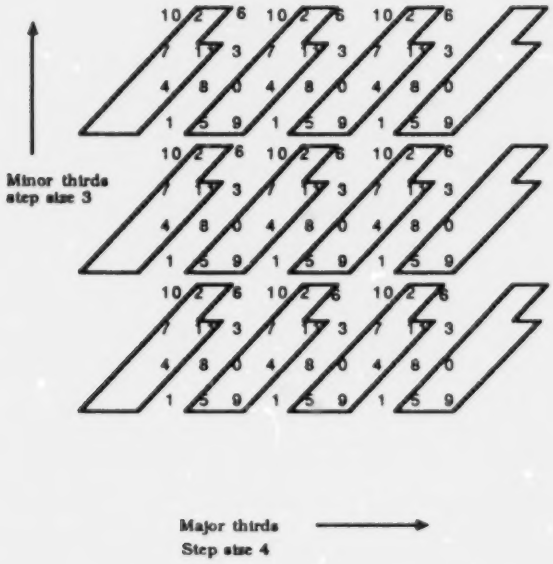


Fig 2 Harmony Space with semitone number labelling

The second point of detail is as follows. The task had been put to YB as one of simply reading out the names of labels and naming the direction of window shifts. However, it was explained to YB that this was equivalent to one form of traditional harmonic analysis. He became very interested and wanted to know what harmonic analysis was about. It was explained that this was equivalent to identifying the positions that chords occupied in the key window, noting what shape they were, and noting where the key window moved. YB then asked (referring to ideas mentioned

during the five minute introduction) if the shape did not follow automatically from the key window position. It was explained that they did normally, but that the key-conforming chord shapes might be overidden at any time in ordinary musical usage. The problem of tracking modulations was to determine whether departures from default shapes were isolated or were most economically described in terms of a movement of the key window. This explanation appeared to interest YB and he said that it made sense to him, although he was unsure how it fitted into the wider context of musical practice.

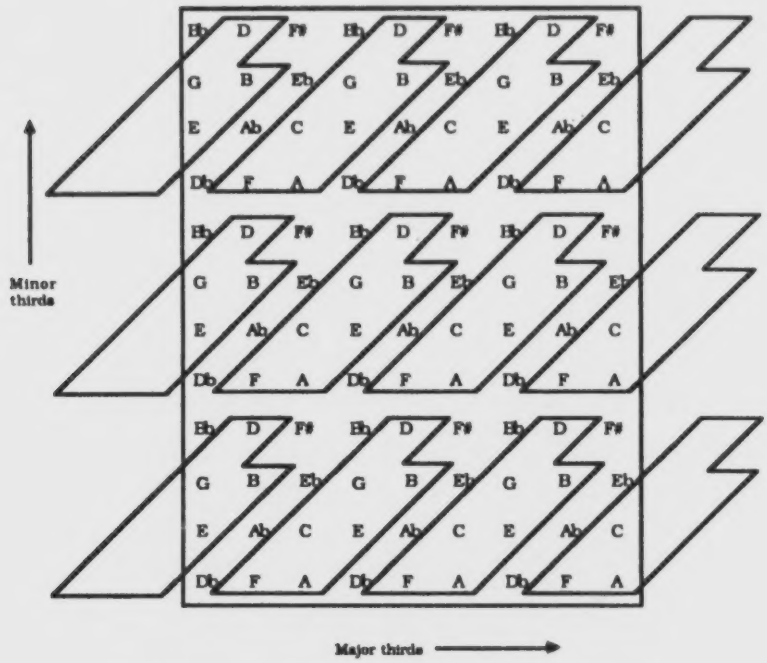


Fig 3 Harmony Space with alphabet name labelling

4 Session 2: BN:
an eight year old's experience with Harmony Space

BN is an eight year old interested in music, but with essentially no musical training. He can play a single one-finger tune on keyboards, taught to him by a friend, but that is all. BN does not have any musical instruments in his house or taught to him at school. BN had a single two hour informal session with the makeshift prototype of Harmony Space. As in all of the remaining sessions, this session was videotaped

and audio-taped, but only the Harmony Space window was videoed (otherwise the most important information for most purposes - i.e the patterns on the screen would have been too small to distinguish). BN was given demonstrations and explanations and asked to perform various tasks. The session was informal, and the investigator taught BN in an opportunistic way, using a variety of approaches, trying to illustrate as many of the potential activities as possible, continuously reassessing what to try next in the light of what BN seemed able to do. BN was enthusiastic about the interface and the (very long) session, wanting to carry on longer. Rather than report chronologically on this long and meandering (though highly energetic) session we will pick out the most interesting points.

Playing chord sequences

Towards the end of the session. BN was taught to play the chord progression of Sam Cooke's "What a Wonderful World":

I VI IV V I

which he learned in less than a minute, after being shown it twice. This was not too hard for him, since by that point he had more or less mastered the three chord trick after many demonstrations, though with a lot of difficulty. The main problem was that at first, he seemed to find it hard to differentiate the different note circle positions within the key window. As the session wore on, BN seemed to get more sense of place in the key window, though he never seemed to become entirely confident of remembering where "home " was. Initially, for example BN found it hard to play a chord progressions such as I IV V I when it was demonstrated, although towards the end of the session he could do this with relative ease.

The 'tritone substitution' game

A game called "tritone substitution" was invented during the session. BN was eventually able to make accurate tritone substitutions, with some help when he made slips, through playing this game. For this game, the default quality of chords with chromatic roots was set by the instructor to dominant sevenths, in accordance with one widespread jazz practice with chromatic chords. The essence of the game, in the version played with BN was that the demonstrator would play chord progressions of

varying lengths, that were cycles of fifths ending on the major tonic. BN was asked to count the number of steps, then play a chord progression of the same length along the chromatic axis to the tonic. (More sophisticated versions of this game were invented in later sessions.) The two progressions are of course aurally and functionally very similar in certain respects. BN's way of carrying out the task was to count as the instructor made his move, then count *away* from the tonic along the chromatic axis, finally playing chords on the return journey. An example move in the 'game' was;

Demonstrator VI II V I
BN IIIb IIdom7 IIbdom7 I

Distinguishing chord qualities, visually and aurally

The demonstrator asked BN at various points whether the various chords with three notes in them (triads) had different shapes or the same shape. At first, BN was unable to distinguish between the different triad shapes from memory (the investigator forgot to make use of the option to hold examples of the different shapes in different places on the screen, which would have simplified this). BN seemed to say in response to questions at various times that each degree of the scale had either the same shape chord, or that each had a different shape. However, as the session progressed, BN became able to distinguish between major, minor and diminished triads by shape. One of the most interesting points in the session was when BN spontaneously announced excitedly that isolated chords seemed to be 'saying' different things, apparently being particularly struck by the sound of a minor seventh, and claiming that some of the chord progressions as a whole seemed to be saying things like "I love you". It may be that BN was starting to distinguish chord quality aurally without being explicitly taught to do so, but it is hard to be certain. It might be interesting for further investigations to concentrate on the development of chord quality differentiation.

Harmonic analysis

BN also tried doing some functional harmonic analysis. Much as in YB's session, this was billed to him as a game in which the demonstrator would make a "move"

that involved either playing a chord or moving the key window, after which BN would announce the name of the chord played or the direction of movement of the key window. As the root names on the prototype are not labelled, BN was given a labelled diagram (fig 1) to refer to. BN had quite lot of difficulty mentally transferring names from the sheet to the screen, but got better at it. It was also very difficult for BN to identify the direction in which the key window moved. (It can be difficult for adults as well, since the routines for moving the key window actually erase and redraw it, taking about half a second.) It would be far easier to carry out this task if the key window visibly slid through intermediate positions.

Understanding the repeating nature of the key-window pattern

After initial confusion, BN had no difficulty realising that equivalent positions in the repeated pattern were functionally equivalent. He would happily move to a different area of the screen to continue the same task if a bug caused a collection of note dots to remain on after they had finished sounding, cluttering up the display (which sometimes happens on the prototype).

Making straight line gestures

BN had some difficulty initially in making straight lines along the semitone axis, veering off upwards or downwards, perhaps confused by the diagonal patterns of the key window. BN had some problems when asked to miss out notes outside the key window (chromatic notes), although he improved at this. (It can be difficult for adults to pick out this diagonal unerringly too. In informal paper and pencil versions of this task, it appears to be easier to track straight lines against the background of the Longuet-Higgins rectilinear key windows than against the background of the diagonally sheared Balzano key-windows - a possible argument in favour of the rectilinear Longuet-Higgins' based configuration).) BN's difficulty underlines the fact that it might be helpful to add a 'constrain' key to the design, along the lines of the 'constrain' key in Macintosh drawing programs. Depressing this key forces lines being drawn to conform to the nearest of the eight principal compass bearings. BN found it difficult to play cycles of fifths, often straying off the path, though he learned to do this eventually. He seemed to find this easier than following the SE semitone and scalar axis.

Establishing a common fund of example pieces

One problem was finding songs that we both knew. Eventually we worked with "Hey Hey we're the Monkees", Phil Collins' "In the air tonight", Stevie Wonder's "Master Blaster", and Sam Cooke's "What A Wonderful World", all of which he more or less knew. BN seemed to find "Master Blaster" particularly interesting, as we played it as a scalar progression followed by a (dorian) three chord trick on the fifths axis.

Aeolian I VII VI V IV I VII I

Following simple, general, compositional strategies

Towards the end of the session, BN had no real problem "composing" and playing chord progressions that consisted of cycles of fifths of various lengths to the tonic. These were billed to him as 'routes down the diagonal towards home'. For example, BN came up with the following chord sequence from following this general instruction

VI II V I

BN was interested to learn that Stevie Wonder (1976) had come up with this chord sequence independently ("Isn't She Lovely") We also discussed the idea that more or less any chord progression would 'do' that started on I and ended on I. BN was encouraged to make up his own chord progressions and spontaneously suggested some variants on the diatonic cycle of thirds, which does not seem to have been shown to him.

Black and white notes

BN asked about the meaning of the black and white areas and this was explained to him on the basis of black and white notes on the piano.

Practical problems

Some of the major problems encountered were as follows. Firstly, the cursor in the

makeshift prototype is difficult to manoeuvre even for a computer-familiar adult, because of its treacle-like response. BN had extra difficulty, as he did not properly learn to pick up the mouse and replace it when he ran out of space on the desk. Secondly, the registration between graphic and sensor areas is slightly out of alignment on the prototype, so without precise (but systematically off-centre) playing, the wrong chord often lights up and plays. We will call this 'mis-triggering'. BN taught himself to compensate for this during the session, though getting the position just right in the face of the treacle-like response was hard physical work.

5 Session 3: RJ: virtuoso visual performance

RJ, a 24 year-old research student with a psychology background, had a single session lasting 1 hour and 20 minutes. RJ has a lot of experience with computers and interfaces. He has no previous knowledge of the project beyond a familiarity with its general aims. RJ has had more or less no previous musical experience that made any impression. RJ abandoned an attempt to learn to play the trombone as a child, being unable to produce any notes. He learned to play a few tunes on a recorder at school by rote at an elementary level, but has long since lost this ability. RJ has never sung (he wasn't allowed to join the school choir) or played any other instruments (except an attempt at drums). RJ has no knowledge of music theory. He never got as far, for example, as finding out what sharp and flat signs mean.

Playing and recognising chord sequences: finding landmarks

RJ learned to "play" harmony space very rapidly, and had no difficulty in remembering the position of the major and minor tonal centres throughout the session. He could name these as such. RJ found it easy (apart from mis-triggerings due to the bad alignment of the interface) after an initial demonstration to play 'a cycle of fifths from major tonic to major tonic', (or minor tonic to minor tonic) when requested in these words. He did this several times on request at different points in the session, producing for example;

C ma7 F ma7 B half diminished 7 E mi7 A mi7 D mi7 G dom7 C ma7
(a cycle of fifths from major tonic C to major tonic C)

A mi7 D mi7 G dom7 C ma7 F ma7 B half diminished 7 E mi7 A mi7
(a cycle of fifths from minor tonic A to minor tonic A)

He could also play such progressions consistently on request diatonically or chromatically, after being explained the distinction. For example,

C ma7 F ma7 Bb dom7 Eb dom7 Ab dom7 D mi7 F# dom 7
B half diminished 7 E mi7 A mi7 D mi7 G dom7 C ma7
(chromatic cycle of fifths from a tonic major C to a tonic major C)

RJ also found it easy to play chromatic and scalic progressions on request, for example

Bb major C minor D minor Eb major .... etcetera
(scalic progression in Bb major)
Bb major B major C minor C# major D minor
(chromatic progression in Bb major with chromatic chords[4] set to major)

though he complained, with some justice, on discovering that progressions *up* the scale moved *down* the screen that this was "bad interface design" (though of course, there is no simple way around this without disturbing the other axes). RJ could identify a sequence played by the demonstrator in terms of whether it was diatonic or chromatic, and whether it moved on the cycle of fifths or scalar axis. RJ could distinguish major, minor and diminished triads on sight, and was able to reproduce the rule for "growing" chords easily. He built up a small vocabulary of progressions he could play easily in the session, namely the three chord trick, the cycle of fifths version of "return home", and simple oscillating modal harmonic ostinati. To give just one example of each kind of progression:

I IV V I
I VII II VI II V I
(Aeolian) I VII VI VII I

[4] In this context, the term "chromatic chords' is used to refer to chords with chromatic roots.

RJ learned to play many of these progressions from abstract specifications, rather than by demonstration. This is illustrated by the fact that, when performing diatonic cycles of fifths, RJ used a different route for the tritone leap than the author habitually uses.

Tritone substitution - a more sophisticated version

Tritone substitution was explained to RJ, in more general terms than to BN, and he was invited to make tritone substitutions of a length chosen by him, starting from a point chosen by him within the sequence and ending on I in the following chord sequence.

VII III VI II V I

RJ's modified version was

VII III VI II IIb I

More than an hour after the investigation, in response to a question about the length of the tritone substitution route he had made, RJ was able to describe his chromatic route correctly as follows - "a short one - just one black note".

Harmonic analysis

RJ was able to carry out the functional harmonic analysis task easily on an improvised chord progression.

Composing chord sequences

When invited to make up some interesting chord sequences, RJ tried several different chord arities under his own initiative, particularly liking fifth diads, and complaining that some of the higher arities were "too jazzy". RJ composed the following chord sequence.

IIb I IV V I

RJ was able to remember the order of chords in this progression an hour and a half after the session. RJ also spontaneously invented the idea of a tritone substitution

chromatically moving scalically *upwards* to the tonic in the tritone substitution game, by taking a 'left turn' instead of a 'right turn'.

Accompanying existing songs

RJ played the chord progressions correctly (apart from occasional mis-triggerings) after brief demonstrations or abstract specifications, was able to accompany the demonstrator singing the choruses of Randy Crawford's "Street Life", Rogers and Hart's "The lady is a tramp", Phil Collins "In the air tonight" and Sam Cooke's "What a Wonderful World". This was not a self-indulgence on the part of the investigator; the role of singing in Harmony Space (or any other) musical instruction can play a vital role which will be discussed in the report on the next and last session.

A visual approach

The investigator tried to illustrate almost every claim in Part II in some form in this session, with little regard for RJ's wish to rest and go to lunch. This was too much to try to cover with a beginner in less than 90 minutes, and RJ complained of 'severe brain overload'. RJ said afterwards "Well, it was very interesting [..] cos y'know I couldn't do that on a keyboard [..] the nice thing is, you do see that it's simple." However, RJ said later (in contrast to BN and LA (the last subject), who both had longer sessions) that he approached the tasks visually and claimed that he hadn't absorbed the relationship between the visual and auditory patterns very much.

6 Session 4 HS: Analysing Mozart's Ave verum

The next subject, HS, is a 37 year old from an Indian background who studied at an Indian school in Kenya. HS has never learned to play an instrument or studied music: either Western or Indian. She sang Indian music in her class at school. She has been exposed to Western pop music since school days. HS works professionally as a secretary. The session with HS lasted just 30 minutes.

HS was taught by demonstration and instruction how to play I IV I V I chord sequences in the major and the minor. It was attempted also to teach HS slight variants of this chord sequence, namely the chord sequence for Transvision Vamp's

'Baby I don't care' (I IV V I) and the Archies' 'Sugar Sugar' (I IV I IV I IV V I) HS found it hard to remember the two different orderings, often confusing one with the other. Problems were greatly compounded by the mis-triggering problems discussed in the other sessions. There was no time to practise these to sort out the problems. HS was reasonably well able to remember the location of the major tonic but found it hard to remember the location of the minor tonic.

After an initial introduction to the interface of five or ten minutes, HS was asked to produce a functional harmonic analysis of a reduced version of the chord sequence of the first part of Mozart's Ave Verum. This was done as follows. The demonstrator used Harmony Space to play a reduced version of the chord sequence (as triads, in root inversion, in close position), with the chord sequence taken to be that given below. The key window was moved to the dominant just before the first II#3.

I II V I V I VII I V
V II#3 III II#3 V I II#3 V

HS was asked to identify the functional names (I, VII, etcetera) of the chords as they were played, by means of the labelled diagrams, and asked to identify the direction of the modulation when it occurred. HS produced the following analysis verbally, chord by chord.

I II V I V I VII I V
(key window moves on upward right diagonal)
V VI V I IV V I

There was little effort to explain what the task meant, due to time constraints. Unlike RJ, who was taught to say things like "modulate to V", by looking up labels for points of the compass in terms of neighbours of I in fig 1 (e.g. NE = V E = III S = VI and so on), HS was not taught functional names for the modulations, being asked only to identify key window movement in terms of points of the compass.

7 Session 5: LA: putting it all together and making sense of the sounds

LA is a French research student of age 29 with a computing background. LA had no knowledge of the project before the session, other than a general familiarity with its goals. To all intents and purposes, she has no formal knowledge of music or instrumental skills. She took part in class sol-fege singing at primary school and listens to pop music, but she has no theoretical knowledge about sol-fa. She learned very elementary recorder playing by rote at about age eight, but could only vaguely remember that she learned to read white notes and black notes and rests; she cannot now remember what the various signs mean. She is unaware of the meaning of sharp and flat signs.

Playing and inventing chord sequences

This was the most relaxed and least pressured session, since a little more time was available. The session was split into two halves with a break for LA's evening meal. The first session lasted nearly an hour, the second session one hour and forty minutes. In the first session, LA learned to play the major and minor three chord tricks, in the following forms:

I IV V I
vi ii iii vi

LA learned that they could be played in different keys. She accompanied the demonstrator, who sang the 'Banana Split Show' theme song and Steve Miller's 'Abracadabra'. LA played both chord sequences correctly, thus;

I IV I V I
vi ii iii vi

LA found it easy to differentiate the major, minor and diminished shapes but did not always remember their names. She found the traditional vocabulary for chord size ('triads, sevenths, ninths', etcetera) confusing and hard to remember, but found the chord-arity terms ('one, two, three, four') to be straightforward. LA learned the

strategy for playing the "return home" plan (major or minor) in the cycle of fifths case. We discussed the fact that Cannel and Marx call this progression the " classical pattern" and that it is widespread in Western tonal music. She made up her own instantiations of the return home plan, for example:

VI II V I

Discussing music theory

LA was concerned to find out the difference between the white notes and the black notes. She was given the same explanation as BN: that they correspond to the white notes and black notes of the piano, and that for many musical purposes, predominantly only the white notes are used. The investigator also explained the fact that the set of privileged notes can change in the middle of a piece of music. It was pointed out that only in Harmony Space are such changes reflected by a shift in the black/white coding (there was a keyboard nearby to point to, although we did not play it).

LA clearly wanted more theoretical background, so fig 2 was used to explain the arrangement of notes in semitone height. It was explained that there were only 12 differentiable notes in an octave in the western pitch set, that octave equivalence allowed these to stand for notes in whatever octave, and that notes on the *semitone axis* of Harmony Space were arranged in order of pitch, both numerically and aurally. To clarify things further, the circles were set to play single notes, and it was demonstrated that their pitch remained invariant when the key window changed position. (To be precise, the octave register of the notes sometimes changed, since the current major tonic is always assigned the lowest pitch, with the other 11 notes in an ascending chromatic scale above it, which was explained to LA. The investigator was worried that this extra detail would confuse LA, but she seemed happy with this explanation. An extra circle display such as discussed in section 3.3 of Chapter 7 might well facilitate this kind of explanation.)

Practical use of the arguments for the centrality of the tonic

Early on in the session, LA seemed to learn the rule for constructing scale-tone chords very easily, and on request could show how to construct triads, sevenths and ninth chords for different degrees of the scale, irrespective of key. LA was taught the locations of the major and minor tonics not just by demonstration but by the 'centrality' arguments of part II. Later in the session, when asked to show the locations of minor doh, LA appeared to make use of both the chord construction rule and of the 'centrality argument', tracing out imaginary minor triads with her finger on the screen, looking for the middle one. However, later in the session, LA was not able to remember the growing rule at one point (although she knew the *sort* of rule) and had to be reminded.

Transfers

After that, it was time for the meal break. After the break, LA learned to play both diatonic or chromatic cycles of fifths on request, starting from the major or minor tonal areas (sometimes she got confused about the location of the minor tonal centre, but mostly had a good picture of where it was).

When she "ran out of screen" on the interface when moving in a straight line towards some pre-specified functional destination, she was taught to transfer to an equivalent position in the pattern, though she sometimes made mistakes on identifying an equivalent position. LA called these "transfers" (they didn't previously have a name). As a result of this instruction, and perhaps partly as a result of the fact it was explained verbally, rather than being demonstrated, LA came up with a slightly different way of playing diatonic cycles of fifths than either RJ or the author. LA picked a representative of VII downwards and to the left where RJ picked a representative downwards and to the right (it is interesting to note that LA is left-handed.)

The semitone axis

LA learned to play progressions up and down down the semitone axis. She found it very easy to play the progression for Phil Collins' "In the Air Tonight"

(Aeolian) I VI VII I.

At first, given an abstract specification in place of a demonstration, LA tended to play

(Aeolian) I VI I

LA found it harder to play Dorian modal harmonic ostinati than their Aeolian counterparts - though she could play either. LA accompanied the investigator singing Michael Jackson's "Billie Jean"

(Dorian) I II III II I

Tritone substitution

LA was able to watch the demonstrator play a "return home chord sequence" and modify it using a tritone substitution; like BN, she found it easier to identify the tonic she was working *towards* and then work backwards rather than to work forwards.

Composing musically 'sensible' chord sequences with a modulation

LA also learned to play a version of Return Home that includes a modulation. This was billed to her as a special case of the general strategy that chord progressions should start and finish on the tonic. The instructions for the game are as follows. First of all one decides whether to play diatonically or chromatically, and major or minor. Then, one plays a home chord, chooses some other chord and begins a progression down the cycle of fifths axis from that chord back to home. One is allowed to make an arbitrary modulation at any point during the journey home. After the modulation, the same trajectory should be continued, observing key constraints if appropriate, to the new home. This game is quite complicated for a beginner, because it is necessary to realise that the trajectory should keep moving regularly with respect to the *underlying note grid*, and *not* relative to the functional positions (that shift with the key window). Also, one must learn to avoid progressions with 'pointless' initial jumps and modulations, for example progressions which reach the tonic by the very act of modulation. LA played the game using chromatic cycles of fifths rather than diatonic cycles of fifths. This game can very easily lead to long

progressions, and typically involved 'transfers', but after some practice, LA definitely got the hang of the game and could play the major or minor version of the games (making new choices of modulation point and direction each time) more or less reliably on request.

Analysing Mozart's Ave Verum with apparent comprehension of the task

LA analysed the Mozart "Ave Verum Corpus" and appeared to understand fairly clearly what the task was about. The task was billed to her as being to identify chords by location in the key window, to identify when they did not follow the "normal construction rule" and to announce modulations in forms of words such as "modulation to V " etcetera. She was asked to use Fig 1 to translate directions of key window move into this form, effectively using chord function names as labels to describe points of the compass. At first LA found it almost impossible to recognise which way the key windows were moving, but after we arranged the key window field so that the "edge" of the field was visible, the task became much easier. LA insisted that the task made good sense to her.

Putting together two chord progressions

LA learned to play the entire chord progression to Randy Crawford's "Street Life", (verse followed immediately by chorus then all repeated). The task was given to her in the following terms: play the minor three chord trick followed by a diatonic progression down the fifths axis from minor tonic to minor tonic. LA made some slips at first but could soon play all of this quite positively.

(minor) I IV V I IV VII III VI II V I

As it happens, the chord progression was played in triads (not sevenths). The harmonic minor in the chorus was not introduced, as this seemed to be too much to deal with all at once.

The role of singing : fitting together melody and chords

LA accompanied the demonstrator/investigator who sang while she played. Singing by the demonstrator was deliberately used to accompany chord sequences whenever

possible. Jazz players often claim that the very best advice in learning to play is to 'sing while you play' (Sudnow, 1978). All of the students were encouraged to sing, and all claimed to be too shy. But singing by both the investigator and student seemed to play important roles in LA's session. In earlier parts of the investigation, LA had complained quite strongly that she could not see how the investigator's singing of "Abracadabra" and the chords she (or he) were playing fitted together. This led to the investigator asking her to count as she played rhythmically, in order to make the chords and singing synchronise as well as possible. This was very difficult given the treacle-like cursor action. The investigator also tried playing Harmony Space himself while singing with as regular a delivery as possible in order to resolve this point. Once LA got the hang of playing with regular counts of four, (as far as the cursor action permitted), LA was much happier and agreed that she could sense the connection a little better, but said that the melody and chords still seemed somewhat unconnected.

About half way through the second session, LA suddenly announced that the singing and the chords were starting to fit together. she also started spontaneously singing to herself as she played, often singing roots. She also announced that she could hear how the different shapes had different sounds.

LA's reaction

LA was due to go to a party half way through the second session, but kept two friends waiting for one hour despite protests from her friends while we finished the investigation. This was not due to the novelty of playing with a new interface - LA has played with very many computer interfaces. She said that the time flew very quickly and that she was really enjoying herself and learning a lot. She said such things as "I can see now it is very clear in my head", "I think I learned a lot", "Now I can see the pattern", as well as going over specific detail of what she had done. LA was sufficiently confident after the session to be convinced that the skills would transfer easily to the piano after, say, half a dozen sessions, even on the existing implementation of Harmony Space. Needless to say, this speculation remains open to empirical investigation, but it is noteworthy that the session produced this level of confidence in a novice.

8 Conclusions

The investigation was a qualitative evaluation, based on single sessions with 5 subjects. In practical terms, it has been demonstrated that musical novices from a wide range of cultural backgrounds and ages can be taught, using Harmony Space, in the space of between 10 minutes and two and a half hours to carry out tasks such as the following :

- Perform harmonic analyses of the chord functions and modulations (ignoring inversions) of such pieces as Mozart's "Ave Verum Corpus". The harmonic analysis was performed on a version played in triads, in close position, in root inversion. This task was carried out by one subject after only 10 minutes training.

- Accompany sung performances of songs, playing the correct chords, on the basis of simple verbal instructions or demonstrations. Songs were selected with a range of contrasting harmonic constructions.

- Execute simple strategies for composing chord sequences such as the following 'musical plans'; return home, cautious exploration, modal harmonic ostinato and return home with a modulation.

- Modify existing chord sequences in musically 'sensible' ways, for example performing tritone substitutions on jazz chord sequences.

- Play and recognise various classes of abstractly described, musically useful chord sequences in various keys both diatonically and chromatically.

- Carry out various musical tasks, such as to recognise and distinguish chord qualities; to use the rule for scale tone chord construction; to locate major and minor tonic degrees in any key; and to make use of the rationale for the centrality of the major and minor tonics.

In all of the sessions, bearing in mind the cultural diversity (China, France,

Kenya/India) and range of ages in the sample, popular music proved an invaluable lingua franca. The fact that the investigation made use of popular music proved useful in a number of ways: all of the subjects had heard some rock music and seemed motivated to play it. Apart from YB's very unusual case, it was mostly easy to find musical examples familiar to the subjects to illustrate the theoretical points. It is unclear how well classical music would have fulfilled the same role.

Experience with Harmony Space bore out the jazz adage that singing while you play can be beneficial to learning (Sudnow, 1978).

Although it was clear from expressivity arguments alone that complex musical tasks could be carried out mechanically in Harmony Space by people who can draw straight lines, etcetera, it was unclear before the investigation whether an *intuitive* 'feel' for the tasks would be fostered by the treacly, clumsy implementation. Surprisingly, on the evidence of the session, at least one of the subjects appears to have started to develop such an intuitive 'feel' for some of the tasks.

Conclusions about the interface

• It has been demonstrated that non-musically trained users (including a child) can use and make sense of the interface.

• The power of Harmony Space has been demonstrated for teaching and performing simple harmonic analysis. The analysis task has unquestionably become easy to do, whether mechanically or with understanding. It is unclear by what other parsimonious mechanism, useful for a wide range of other musical tasks, complete novices could produce similar analyses after five minutes training.

• In harmonic analyses, we have ignored inversions (i.e. octave height information) and concentrated on identifying roots, chord qualities and shifts of tonal centres. With a variant versions of the interface, where octave height information (i.e. Longuet-Higgins' third dimension [5]) is restored using colours or shading, the full

[5] In terms of Balzano's theory, the z axis may be seen in much the same way: the set of equivalence classes for C_{12} stacked along C_{∞} arranged in order of octave height.

task could be carried out by comparing shadings to derive inversion information, without changing the interface in any other way. For example, if octave height were to be mapped so that 'darker = lower', the 'place' of the darkest note in a chord (in terms of the scale-tone chord-construction rule) would correspond to its inversion.

• Using a musical terminology that re-interprets traditional terms according to how they appear in Harmony Space i.e. "chord arity" may allow musically untrained novices to acquire a vocabulary for articulating intuitions about Harmony faster, and may help them retain the vocabulary longer. It must ultimately depend on the purposes of the learner whether it is preferable to learn the traditional vocabulary for Harmony, or the Harmony space neologisms on initial exposure to Harmony Space.

• Harmony Space is more useful than, for example, a piano for enabling musically untrained novices to perform the tasks we have considered. As a thought experiment, let us imagine a process of analysis, similar to YB's, being attempted by a novice watching a slow performance in root inversion, in close position, on a piano. If the keys were labelled, it appears vaguely plausible at first that the task could be done, until we remember that the labels would need to move when the pianist modulated. This could be remedied using back-projection and transparent keys, but would be fatally inadequate for control (as opposed to recognition) tasks, since, for example, chords are hard and irregular for a novice to play in remote keys on a piano. A modulating piano (a piano with a special lever for retuning the entire piano up and down a given number of semitones) or its electronic equivalent, at first appears to get around this problem, but it is unclear how such an instrument could go very far to help novices perform most of other tasks discussed in this chapter (e.g. learning to play complex, varied chord sequences in a matter of minutes, learning to make tritone substitutions intelligently, distinguishing chord qualities irrespective of the degree of scale of the root, composing new chord sequences in a purposive manner, playing and recognising regular progressions while modulating, judging the harmonic closeness of modulations and so forth). These problems could no doubt be addressed in a piecemeal fashion, but we would be in danger of re-inventing harmony space: Balzano's uniqueness proof establishes that harmonic relationships are inherently two-dimensional , not just linear or circular.

All of the above listed tasks, and many others, can be performed using the

parsimonious Harmony Space control/display grid *as it stands*. The parsimony of the Harmony Space interface means that each different use of the interface for a fresh task has a chance to imbue the same few locations, shapes and movements with a fresh, but consistent, layer of musical meaning. The irregularity of the diatonic scale, noted by Shepard (1982) at the head of part II is reflected faithfully in the shape of the irregular key window shape of Harmony Space. This means that every location and shape in Harmony Space can come to have unique musical associations for users (similar to those that chord symbols and CMN symbols can come to have for traditionally trained musicians). The investigation suggests that this kind of musical meaning can start to emerge for novices in a matter of minutes to hours, even on a poor implementation.

• Harmony Space's group theoretical principles, and Balzano's uniqueness arguments, give us general grounds for believing that Harmony space is particularly well suited for Harmonic analysis, whether by mechanical methods or with understanding.

• Despite the mechanical style of presentation of the task of harmonic analysis to YB, his performance, and his query about the meaning of the task illustrate the fact that Harmony Space is not only well suited to the performance of harmonic analysis, but also well suited to support explanations to novices of the meaning of the process.

9 Extensions

A straightforward extension of this study would be to take a small number of students, (perhaps some in pairs) for half a dozen weekly sessions and then find out the degree to which the composition, analysis, accompaniment and theoretical skills learned could be transferred to a keyboard.

A useful part of such an investigation would be to invent more 'games' to play and to experiment with tasks to carry out in Harmony Space.

A useful contribution to such a study would be to make a version of the key window display available in which the key windows are graphically 'joined up', to see if this improves or hinders development of a sense of location in the key window.

A simple implementation project would be to implement the shading of notes on the display according to octave height, and the use of mouse buttons to control inversion. This would be a logical extension to the Harmony Space design, since it simply corresponds to making Longuet-Higgins's z-axis visible and controllable. Amongst other things, this would permit experiments with novices on the production of harmonic analyses that took into account inversion

10 Further work

An interesting research project in the light of the present research would be to implement Harmony Space in a rapidly-responding, accurate, properly-labelled form, with DM secondary functions, smoothly sliding key windows and control and display of inversions. This would be more suitable for practical experimentation, for more refined harmonic analysis and might well help develop a better gestalt/intuitive feel for harmony. Since important proximity relationships are preserved in the auditory/haptic/visual mapping, we hope that a fast, accurate implementation of Harmony Space would engage perceptual gestalt grouping mechanisms in a similar way all three modalities, and that intuitions would start to 'seep into the fingers, ears and eyes' of some users, almost independently of what they *thought* about what they were doing.

An interesting research question, that might be investigated in the light of the present research, would be to find out whether interfaces based on (ubiquitous) group theoretic characterisations of other domains might be designed that translated proximity relationships that happen to be hard to intuit in the abstract into haptic [6], visual, easily assimilable form.

[6] Balzano (1982) notes that the relationship of group theory to perception is close, and has been investigated by (Cassirer, 1944)

Chapter 10: Conclusions and further work

> *"..one must learn to argue with unexplained terms and to use sentences for which no clear rules of usage are yet available."*
>
> Paul Feyerabend p 256 (1975).

In this final chapter, we will bring together the most important of the conclusions from throughout the thesis. We will also draw several new conclusions.

1 Contribution

1.1 Overview of contribution

This research has produced a theoretically-motivated design for a computer-based tool for exploring tonal harmony. A prototype version of the interface has been constructed to demonstrate the feasibility and coherence of the design. A qualitative investigation has demonstrated that the interface can enable musical novices from diverse backgrounds to perform a range of musical tasks that would normally be difficult for them to do. In addition, the research has contributed towards the design of a knowledge-based tutoring system to assist novices in composing chord sequences. The approach is built on the exploration and transformation of case studies described in terms of chunks, styles and plans. Programs to demonstrate the feasibility of key parts of the knowledge-based approach have been implemented.

As an organisational device, these contributions have been divided up into three categories: contributions to Human Computer Interface research, contributions to Artificial Intelligence and Education, and contributions to Music Education. However, many of the contributions belong under more than one heading.

1.2 Contribution to Human Computer Interface research

1.2.1 Design of an expressive interface

An interface has been devised which makes use of representations derived from two cognitive theories of harmony. Design decisions have been discussed which employ the representations in the interface in ways which give the interface particularly desirable properties. The properties are given below. (Three of the most useful traditional 'interfaces' are referred to in the next few points for comparison, namely; common music notation, the guitar fretboard and the piano keyboard.)

1 Chord progressions fundamental to western tonal harmony (such as cycles of fifths and scalic progressions) are translated into simple straight line movements along different axes. This renders fundamental chord progressions easy to recognise and control in any key using very simple gestures. (These progressions can be hard for novices to control and recognise by conventional means.)

2 The same property enables particularly important progressions (such as cycles of fifths) to be manipulated and recognised as *unitary "chunks"*. A common problem for novice musicians is that they only treat chord progressions chord by chord (Cork, 1988).

3 Chords that are harmonically *close* are made visually and haptically close (fig 3 chapter 5). This renders other fundamental chord progressions (such as IV V I and progression in thirds) easy to recognise and play. (On a piano, for example, these progressions require gestural 'leaps' to be made that can be particularly hard to judge in 'remote' keys. All keys are rendered equally accessible in Harmony Space.)

4 Similarly, *keys* that are harmonically close are made visually and haptically close. This makes common modulations easy to recognise and play. (In common music notation, or on either traditional instrument , considerable experience or knowledge can be required to work out key relationships.)

5 The basic materials of tonal harmony (triads) are made visually recognisable as maximally compact objects with three distinguishable elements. (In other representations mentioned above, the selection of triads

as basic materials may seem arbitrary.)

6 Major and minor chords are made easy to distinguish visually. (This is not the case, for example, in common music notation or on a piano keyboard)

7 Scale-tone chord construction is made such that it can be seen to obey a consistent and relatively simple graphic rule in any key. (On a piano keyboard, for example, this is not the case in remote keys, where semitone counting or memorisation of the diatonic pattern for the key is necessary.)

8 Consistent *chord qualities* are translated into consistent shapes, irrespective of key context. (Again, this is not the case, for example, on a piano)

9 The tonic chords of the major and minor keys are spatially central within the set of diatonic notes, whatever the key context. (This relationship is not explicit in common music notation, guitar or piano keyboard)

10 Scale-tone chords appropriate to each degree of the scale can be played (even when there is frequent transient modulation) with no cognitive or manipulative load required to select appropriate chord qualities (This does not hold with guitar or piano).

11 The rule determining the correspondence of chord quality with degree of scale makes sense in terms of a simple, consistent physical containment metaphor.

12 A single consistent spatial metaphor is sufficient to render interval, chord progression, degree of the scale, and modulation. (Not so in the other three points of comparison)

1.2.2 General characteristics from an interface design viewpoint

The interface also has a number of useful properties in terms of more general human-machine interface design criteria and viewpoints, as follows;

Consistency and uniformity of control and representation:

Harmony Space exhibits a high degree of uniformity and consistency in the way that the phenomena it deals with are controlled and represented. This is held to be a useful trait in human-machine interfaces (Smith, Irby, Kimball, Verplank and Harslem 1982) from the point of view of encouraging ease of learning and use. The relevant consistencies and uniformities of harmony space can summarised as follows;

- a single *spatial* metaphor is used to describe and control pitch, interval, and key-relationships,
- a single *movement* metaphor is used to describe melody, chord progression and modulation,
- a single *spatial containment* metaphor is used to describe and distinguish the diatonic scale, scale tone chord qualities and diatonic chord progressions from their chromatic counterparts,
- a single *spatial centrality* metaphor is used to describe major and minor tonal centres,
- chord qualities have *consistent shapes* in any key. The shapes reveal their internal intervallic structure explicitly,
- a single *proximity metaphor* applies to harmonically close intervals, chords and keys.

All of these metaphors can be seen as aspects of a single spatial metaphor.

Reduction of cognitive load

Cognitive load is reduced by the interface because relationships that have to be learned or calculated in other representations (such as common music notation, or notations based on the guitar or piano keyboard) are represented explicitly or uniformly in Harmony Space. Example relationships include;

- the location of the notes of the diatonic scale in different keys,
- the shapes of different chord qualities in different keys,

- the internal intervallic structure of chords on different degrees of the scale,
- the degree of note overlap between sharp and flat keys,
- the structure of the cycle of keys.

Examples abound of the way in which the interface can reduce cognitive load in particular tasks. To pick two more or less at random: when modulating between two keys, the number of notes in common does not have to be calculated using previously memorised knowledge - it can be grasped visually. Again, when inspecting chord progressions in roman numeral notation by conventional means, a considerable amount of implicit knowledge and processing or memorisation is required to recognise regular chord progressions. In harmony space, such progressions are trivially recognisable in a single mental 'chunk'.

Use of existing knowledge and propensities

Unlike most musical instruments, rather than requiring the intensive learning of new skills, Harmony Space exploits skills that most people already have, such as being able to;

- recognise straight lines
- make straight line gestures,
- keep within a marked territory,
- distinguish points of the compass,
- recognise and find objects that are close to each other,

and so forth. The design of the interface maps complex, unfamiliar skills, normally requiring extensive training, into commonplace abilities that most people already have.

A principled multi-media, multi-modality interface

We have shown how harmony space exploits cognitive theories to give principled, uniform, synchronised metaphors in two modalities (visual and kinaesthetic) for harmonic relationships and harmonic processes occuring in a third modality (the auditory).

1.2.3 Design of a tool that gives access to otherwise inaccessible experiences

Harmony Space gives novices access to normally inaccessible experience in two distinct but related senses described below.

Purposive control of normally inaccessible experience by novices

On most existing musical instruments, (and even on most computer-based editors and interfaces), the gestures demanded of novices to allow them to experiment in a purposive way with the materials of tonal music cannot be performed in real time without considerable pre-existing instrumental skill or musical knowledge. These prerequisites can form a considerable barrier for novices, preventing them from ever carrying out meaningful musical experiments. Harmony Space makes the basic materials of harmony easy to recognise and control in a flexible manner using very simple gestures.

Providing a vocabulary to analyse and articulate experience

Articulating intuitions about chord sequences requires a vocabulary of some sort. The traditional vocabulary is difficult to learn. Novices may have strong intuitions about chord sequences, but may have no vocabulary for articulating and discussing the structure of chord sequences. Harmony space gives novices a simple, uniform spatial vocabulary and notation for articulating and discussing such matters. This vocabulary and notation are consistent and uniform across all relevant levels, from the chord constituent level, through the chord progressions level, up to the level of inter-relationships between modulations. Harmony Space need not cut novices off from traditional vocabulary, but to the contrary, could be used as a complement and means of access to more traditional vocabularies.

1.2.4 Qualitative investigation of interface

It has been demonstrated that non-musically trained users (including a child) can use and make sense of the interface to carry out a range of substantial musical tasks.

It was not clear before the investigation whether an *intuitive* 'feel' for the tasks would come across with the limited implementation. Surprisingly, on the evidence

of the session, at least one of the subjects appears to have begun to develop such an intuitive 'feel' for some of the tasks.

1.3 Contribution to music education

1.3.1 A new approach to music education in harmony

It has been demonstrated with a small number of subjects using Harmony Space that musical novices from a wide range of cultural backgrounds and ages can be taught (in the space of between 10 minutes and two and a half hours) to perform tasks such as the following :

- Carry out harmonic analyses of the chord functions and modulations (ignoring inversions) of such pieces as Mozart's "Ave Verum Corpus". The harmonic analyses were done on a version played in triads, in close position, in root inversion. This task was performed by one subject after only 10 minutes training.

- Accompany songs, playing the correct chords, on the basis of simple verbal instructions or demonstrations. Songs were chosen with a range of contrasting harmonic constructions.

- Carry out simple strategies for composing chord sequences such as the following 'musical plans'; return home, cautious exploration, modal harmonic ostinato and return home with a modulation.

- Modify existing chord sequences in musically 'sensible' ways, for example carrying out tritone substitutions on jazz chord sequences.

- Play and recognise various classes of abstractly described, musically useful chord sequences in various keys both diatonically and chromatically.

- Carry out various musical tasks, such as to recognise and distinguish chord qualities; to make use of the rule for scale tone chord construction; to find the major and minor tonic degrees in any key; and to be able to use the rationale for

the centrality of the major and minor tonics.

1.3.2 Development of a new approach to teaching aspects of musical theory

The power of Harmony Space has been illustrated for carrying out simple harmonic analysis. Analysis is relatively simple to do using Harmony Space, whether mechanically or with understanding. It has also been demonstrated that Harmony Space is not only well suited to carrying out harmonic analysis, but also well suited to support explanations to novices of the significance of the process.

1.3.3 A multipurpose tool

The tool has been identified as being particularly relevant to the needs of those with no musical instrument skills, unfamiliar with music theory and unable to read music, but who would like to compose, analyse or articulate intuitions about chord sequences. Without such a tool, it is unclear how members of this group could even begin to perform these tasks without lengthy preliminary training.

Harmony space may be able to give conceptual tools for articulating intuitions about chord sequences not only to novices but also to experts. The graphic linear notation for chord progressions that emerges from Harmony Space when tracking chord sequences makes various aspects of chord progressions visible and explicit that are otherwise hard to articulate in conventional notations. More expert musicians might also find Harmony Space and its variants useful for such tasks as assessing the harmonic resources of different scales, subsets of scales and so forth. The following uses for Harmony Space have been identified;

- musical instrument,
- tool for musical analysis,
- learning tool for simple tonal music theory,
- learning tool for exploring more advanced aspects of harmony,
- discovery learning tool for composing chord sequences,
- notation for chord sequences,
- notation for high-level aspects of chord sequences (apparently not currently representable *perspicaciously* in any other way).

1.4 Contributions to Artificial Intelligence and Education research

In open ended domains, there is no precise goal but only some pre-existing constraints to be satisfied, any of which may be dropped when some other organisational principle becomes more important. To illustrate a way of dealing with such situations, a new framework for knowledge-based tutoring called MC, to be integrated with Harmony Space, has been described. The purpose of the framework is to support the learning of skills in composing chord sequences. Ways have been outlined of characterising useful forms of higher level musical knowledge about chord sequences and using it to transform musical materials.

1.4.1 Characteristics of MC framework

MC has various properties that make it well-suited to this kind of open ended domain, as follows;

- MC offers complementary, conceptual and experiential learning. The Harmony Space component offers opportunities for experiential learning. MC offers opportunities for more abstract, conceptual learning by providing representations of strategic musical knowledge. The knowledge can be used to transform chord sequences in musically intelligent ways, according to flexible specifications made by a user or teaching strategy. Examples can be explored experientially in one component and more abstractly in the other component.

- A key feature of the conception of MC is the employment of a series of multiple modalities, representations and viewpoints. There are three modalities in Harmony Space: the student may hear the notes of a piece, see its trace as an animation in harmony space, or trace it using appropriate muscular movements. Stored pieces may be characterised from three different viewpoints, as a chunked note level description, in terms of styles and in terms of plans. A given song may be chunked in more than one way - corresponding to different analyses of the piece.

- The MC architecture promotes the use of goal-oriented, high-level strategic thinking. If a student begins from an existing piece, the architecture allows

the student to look at the piece as being goal-oriented according to several competing plan descriptions. Alternatively, the student can start from a schema or musical strategy. There is a strong emphasis in the architecture on metacognitive skills. Rather than encouraging a focus on the low level elements of chord sequences, the student is enabled to consider the ends that chord sequences serve and the kinds of high level considerations that can affect their shape.

• It does not matter whether the student 'learns' any of the plans or any of the songs. Successful outcomes are of two types: firstly that the student devises interesting new chord sequences - though the system holds no targets for the exact form these may take, and secondly that the student comes to realise that chord sequences have a strategic dimension, and that various low level means may be employed to serve various high level musical ends.

• When songs are characterised in the form of musically meaningful "chunks", looking at the effect of mutations is more likely to be illuminating for novices than changing individual notes. Chunks based on Harmony Logo primitives, seem compatible with common intuitions about pieces and informal terms used by musicians.

1.4.2 Wider applicability of MC framework

The MC architecture has been shown to have a number of implications for future Intelligent Tutoring Systems research in open-ended domains, as listed below.

• It has been argued that the Johnson-Laird characterisation of creativity, not previously used in an Intelligent Tutoring Systems context, is particularly apt metaphor for use in representing expertise in any open-ended creative domain.

• The use of object level, style level and goal-oriented or strategic levels of description appears to be potentially relevant in any domain which involves the exercise of open-ended thinking (e.g. engineering design and

architectural design).

• The Paul Simon, or intelligent transformation method appears to be a useful, (though necessarily comparatively weak) technique for learning design skills in any domain which involves the exercise of creativity.

2 Criticisms and limitations

In this section, criticisms and limitations of the research are acknowledged. Many of these limitation are taken up in the extensions and further work section.

2.1 Musical Limitations to the research

There are some aspects of harmony that Harmony Space does not represent well, for example, voice-leading, the control and representation of voicing and inversion, and the visualisation and control of harmony in a metrical context. Harmony Space emphasises vertical aspects (in the musical traditional sense) at the expense of linear aspects of harmony. To a large extent this is a built-in limitation of Harmony Space, although Harmony Space can demonstrate some special cases of voice leading rather well (Holland, 1986b). Some other musical limitations of the research are as follows;

- Some musical terms and notations have been used in unusual ways.
- Some musical neologisms have been coined for which traditional terms exist.
- Harmony Space deals only with tonal harmony.
- Harmony Space is not very well suited to dealing with melody and rhythm.
- Inverted harmonic functions have been ignored.
- The musical knowledge in MC is heuristic and informal.
- MC characterisations of chord sequences are sometimes approximate.

2.2 Shortcomings of the implementation of Harmony Space

The current implementation of Harmony Space is an experimental prototype designed to show the coherence and practicality of the design. At the time at which the work was done, the graphics, mouse and midi programming in themselves required considerable effort to undertake. The prototype served its intended purpose, but is

too slow and lacks simple graphic and control refinements to make it easy to use. In particular, it suffers from the following limitations;

- the prototype is slow to respond and garbage collects frequently, hampering rhythmic aspects of playing the interface,
- the slow response of the prototype means that events that should emerge as single gestalts (such as smooth chord progressions and modulations) are broken up,
- the response problems interfere with the illusion of a direct connection between gesture and response, and can cause considerable frustration,
- the secondary functions such as overiding chord quality, switching chord arity, controlling diatonicism/chromaticism and modulating are controlled in the current prototype by function keys as opposed to pointing devices,
- the 'state' that the interface is in, and the choice of states available is invisible, rather than indicated by buttons and menus on the screen,
- it is not possible to view the grid labelled by roman numeral, degrees of the scale, note names etc.,
- output is not provided in Common music notation, 'piano keyboard and dot' animation or roman numeral notation to assist skill transfer (see appendix 2).
- only the Balzano configuration has been implemented.

2.3 Shortcomings in implementation of MC

The MC framework has not been fully implemented. Simple prototypes to demonstrate the feasibility of all of the innovative components required by the framework have been implemented. Song and plan level descriptions have been implemented in a rudimentary form using the prototype Harmony Logo interpreter. Direct links to Harmony Space are not implemented. The Harmony Logo interpreter as implemented uses a cruder version of Harmony Logo syntax than the programs synthesised by the planner.

2.4 Limitations to the practical investigation

The empirical investigation was a qualitative evaluation, not a psychological experiment or educational evaluation. The sample was small. Only single sessions

were used. The harmonic analyses were performed on reduced harmonic versions played in triads, in close position, in root inversion.

3 Extensions

3.1 Extensions to practical investigation

• A simple extension of the empirical investigation would be to take a small group of students for ten or so weekly sessions and find out the extent to which the composition, analysis, accompaniment and music theory skills could be transferred to a keyboard.

• A useful contribution to such an investigation would be to devise more compositional and other musical 'games' to play and tasks to carry out in Harmony Space.

• Another helpful contribution would be to make a 'clone' of Harmony Space with a small adjustment in the key window display. The adjustment would be to graphically 'join up' the key windows. The aim would be to facilitate experiments to see if this improves or hinders development of a sense of location in the key window.

• Studies into the relative ease of use of the two configurations (Longuet-Higgins and Balzano) would be facilitated if a version of Harmony Space was implemented in the Longuet-Higgins configuration.

• A straightforward extension of the present research would be an in-depth educational evaluation of Harmony Space. It is suggested that Laurillard (1988) would form a good basis for such evaluation.

3.2 Harmony Space and harmonic inversions: programming and experiments

A simple but useful programming project would be to implement the shading of notes on the Harmony Space display according to octave height. The implementation of some means of *controlling* inversions and pitch register in Harmony Space would likewise be valuable and relatively easy. These simple projects would make possible experiments with novices using Harmony Space to

carry out harmonic analyses that took inversion into account. Inversion could be controlled manually by means of mouse buttons (or, more ambitiously, a midi glove). It might be interesting at the same time to experiment with simple 'least change' algorithms to automatically choose inversions, for example so that constituent notes move in pitch as little as possible.

3.3 Compiling a comprehensive Harmony Space dictionary of chord sequences

Among the chord sequences so far transcribed into Harmony Space notation, the density of interesting, almost immediately apparent regularities has appeared to be high. However, the number of chord sequences so far been transcribed into Harmony Space notation and analysed has been limited. Consequently, many regularities that might be apparent in this notation may have been overlooked. For example, Johnson-Laird (1989) has pointed out that the chord sequence for John Coltrane's seemingly irregular "Giant Steps" (Coltrane, 1959) exhibits a regular herringbone pattern in Harmony Space. A straightforward extension of the present project, that could be done with paper and pencil (but more quickly and accurately using Harmony Space) would be to transcribe an extensive library of chord sequences, particularly jazz chord sequences, into Harmony Space notation and analyse them in order to taxonomise such regularities. Depending on what regularities were discovered, this might grow into a larger project with implications for mu[illegible]lucation.

4 Further work

4.1 An improved implementation of Harmony Space

In the light of the present research, a very useful larger-scale design and programming project would be to reimplement Harmony Space in a rapidly-responding, accurate, properly-labelled form, configurable in Balzano's or Longuet-Higgins' configuration, with DM secondary functions, smoothly sliding key windows, constraint controls for drawing straight lines, and facilities for control and display of inversions. An implementation of Harmony Space along these lines would be more suitable for practical experimentation, suitable for more refined harmonic analysis, suitable for supporting a wider range of experiments and might well help develop a better gestalt/intuitive feel for harmony. Such a version of

Harmony Space would be involve many small design decisions (e.g. menu layouts); therefore several design-and-evaluate iterations should be allowed for. This project would greatly facilitate many of the studies proposed above.

4.2 Version of Harmony space with metric display features

In order to understand a chord sequence it is often necessary to know about its metrical context (as in the analysis of "All the things you are"). A useful medium-sized research, design, programming and evaluation project could be based around combining a graphic notation for metre similar to that used by Lerdahl and Jackendoff (1983) with a Harmony Space display, to help keep track of how far a performance in step-time or real-time was progressing with respect to a pre-defined metrical grid (fig 1). Once implemented, such a tool could be evaluated to discover the extent to which it promoted harmonic skill and understanding in metrical contexts.

Fig 1 Suggested display to show point in metrical structure of a piece of events in Harmony Space

4.3 Development of Harmony Logo

A useful medium term design, programming and evaluation project would be to further develop Harmony Logo. Harmony Logo is a programming language that moves voices audibly and graphically around in Harmony Space under program control, in a fashion similar to the way the programming language "Turtle Logo" moves graphic turtles around on a computer screen. This potentially allows the structure of long bass lines and chord sequences to be compactly represented in cases where the gestural representation might be awkward to perform or hard to analyse.

A prototype version of Harmony Logo has already been implemented which allows sounding objects to be moved around in Harmony Space under program control by means of a simple Logo-like language. The prototype has a number of major limitations: there is no graphic playback; the current syntax is clumsy and limited; the facilities for rhythmic control are poor, and it is unclear whether the embedding of harmonic trajectories into the interpreter cycle is helpful. A clean version of this language could be a useful educational tool. The outlines of the project would be as follows;

- review the rhythmic aspects of Harmony Logo for possible redesign,
- re-design and re-implement the language cleanly
- provide a graphic interface to Harmony Space,
- Assess Harmony Logo as an educational tool for exploring harmony.

4.4 Improved implementation of key components of MC framework

The following series of projects would constitute a single, sizeable, programme. The first step would be to re-implement the Harmony Logo interpreter, interface it to Harmony Space and build up style and song libraries. The next step would be to reimplement PLANC (the implemented constraint-based planner forming part of MC) using the new dialect of Harmony Logo and interface it to Harmony Logo and Harmony Space. Empirical investigations could then be performed on the re-implemented and mutually interfaced Harmony Space, Harmony Logo and PLANC to assess their usefulness as a teaching aid for the composition of chord sequences.

4.5 Evolving new curricula and teaching methods for Harmony Space

A medium sized educational project (which could be begun with small pilot investigations, useful in their own right) would be develop new curricula and teaching methods to take advantage of Harmony Space's theoretical basis. The point of this project can be seen as follows. Harmony Space breaks new ground educationally in the following ways;

- using a new conceptualisation of the domain treated,
- teaching skills not normally taught (e.g. composition of chord sequences),
- teaching skills to a group without the normal prerequisites (i.e. basic instrumental skills).

Considerable trial and error may be required to gain experience of the best ways of using Harmony Space effectively in education. Part of the project would be to carry out educational experiments to gain the necessary experience.

• Harmony Space is based on theories that *reconceptualise* the domain in question. The reconceptualisations call for and give an opportunity to make radical changes to the curriculum and teaching approaches in this area. A second part of the project would be to teach students harmony experimentally *in terms of* Longuet-Higgins' and Balzano's theories, using Harmony Space and written teaching materials (as opposed to using Harmony Space to teach harmony in traditional terms). This would require developing a curriculum, teaching approaches and written teaching materials appropriate to the new conceptualisation.

• Students with no knowledge of music theory, no musical instrumental playing skills and no music reading skills are not normally taught non-vocal tonal composition skills in any formal way. Consequently, there is little experience of how to teach such students such skills. The composition of chord sequences is not in itself normally taught to anyone as a subject, perhaps because there is a lack of accepted theory of how to do it or teach it. The thirds part of the project would be to use Harmony Space for building up experience of good practice in teaching this group and this skill.

4.6 Graphic Harmonic analysis using Harmony Space

An interesting medium-sized research project would be to experiment in the use of Harmony Space to allow users to perform a version of traditional harmonic analysis in a graphic, intuitive manner. This proposal would involve using Harmony Space to carry out *graphically* tasks performed *algorithmically* by Steedman's (1972) cognitive modelling programs using a variant of the same representation. Harmonic analysis of a piece of music into chord functions and modulations would be carried out by trial-and-error graphical 'best-fit'. This would involve positioning the key window on traces of pieces (polyphonic or homophonic) played back on the harmony space display to find keys which sections of the trace fitted best. In harness with this procedure, the key window would be used to to identify chord function by

looking for familiar chord shapes and identifying graphically the degree of the scale of chord roots by their position within the key window. Routines in Harmony Space to allow the notes to be 'clumped' together on the display in various ways could facilitate this process. It would also be necessary to have facilities for graphic playbacks of chosen sections of chord sequences, but with key window able to move freely (disconnected from control functions) so that the user could try experimental 'best-fits'. From an Intelligent Tutoring Systems viewpoint, this proposal is an interesting demonstration of the possibility of using one and the same theoretical base to design interventionist tutors or to build open-ended theory-based tools.

4.7 Use of Harmony Space for teaching jazz improvisation skills

Cork (1988) relates that a major problem in teaching jazz improvisation is that students treat each chord as an isolated event and tend to have little conception of the direction and thrust of the chord progression at a higher level. Cork has tackled this problem in an ingenious and innovative way by analysing jazz chord sequences into commonly occurring moves ('Lego blocks') with evocative verbal labels such as 'new horizon' and 'downwinder'. Each building block refers to a progression of chords (in some cases including a modulation and in come cases making reference to metrical context). Cork's system has been used with some practical success to help students conceptualise chord sequences at a higher level and make more fluent and assured improvisations. There appear to be a very strong correspondence between Cork's verbally labelled blocks and short segments commonly seen in harmony space traces. From one point of view, Cork's Lego blocks correspond to small 'macros' in Harmony Space. Three pieces of further work are suggested. Firstly, as a simple pencil and paper exercise, it is proposed that a Harmony Space-Harmony Lego dictionary be drawn up to establish the exact nature of the link between the two systems. The second piece of work would involve empirical investigations to find out whether Harmony Space notation, like Cork's Harmony Lego, can help students conceptualise chord sequences at a higher level and make more fluent and assured improvisations. Thirdly, Cork's approach requires that the student can already play an instrument, read jazz chord symbols, and has experience of the basics of tonal harmony. The third proposal is to discover whether it is possible to help teach improvisational jazz singing, using Harmony Space notation, to novices able to sing,

but with no formal or instrumental musical experience.

4.8 Harmony Space, violins and guitars

The fretboard of a violin or a guitar can be viewed as the 12-fold version of Longuet-Higgins space sheared vertically downwards[1]. The following table shows how the four intervals that can form a well-defined co-ordinate system (perfect fifth, semitone, major third and minor third) are remapped into points of the compass when the Longuet-Higgins 12-fold space is sheared vertically downwards.

Interval	LH	LH sheared vertically downwards
p5	N	N
M3	E	SE
m3	SE	Lost: NE axis becomes a tritone axis
m2	NE	E

The guitar has quite different abstract properties from Longuet-Higgins' 12-fold space. We shall give just four important differences. Firstly, the shearing of the space substitutes the major thirds axis for a tritone axis, gives the two spaces quite different configurational properties. Secondly, the fretboard of a guitar is not wide enough (six strings) for the straight line patterns that occur in cycles of fifths to be clear to beginners. Thirdly, the B string on a guitar breaks the regular grid of fourths, destroying the uniformity of chord and progression shapes. (On a violin, this irregularity does not occur, but the fretboard is too narrow (four strings) for the uniformity of chord shapes on the different strings to be at all obvious to beginners.) Fourthly, the key window is not represented explicitly on a guitar. Still, the special relationship between stringed instruments and Harmony Space does open up some interesting possibilities for future research. (The piano key board can similarly be seen as a thin, one-dimensional window on Harmony Space, running down the chromatic diagonal axis.) An interesting, long term project in instrument design

[1] The fifths axis in a guitar is inverted in sense compared with the convention for Longuet-Higgins diagrams, but this is purely a matter of convention. When the Longuet-Higgins space is sheared *vertically* downwards, the minor thirds axis is 'stretched out of existence'. The 'northern' (fifths) axis is unaffected and the 'eastern' (major thirds) and 'north eastern' (chromatic) axes are shifted one 'place' each clockwise through 45^{0}. But the 'south-eastern' (minor thirds) axis is "stretched" out of existence southwards, and a tritone axis "squeezed" into existence from the south.

would be to construct guitars (technology permitting) with the sheared key window displayed on the fretboard in a moveable form, its location controlled by a foot switch to track modulation. An interesting empirical study would be to find out whether this was helpful when learning to improvise over rapidly modulating chord sequences (such as Frank Zappa audition pieces).

We now come on to a series of larger scale proposed research projects.

4.9 Microworlds for other areas of music

An interesting large scale research project in the light of the present research would be to design microworlds based on appropriate theories of aspects of music such as rhythm, melody and voice-leading to complement harmony space. It is suggested that the theories of Longuet-Higgins and Lee (1984) and Lerdahl and Jackendoff (1983); Levitt (1981,1985); and McAdams and Bregman (1985) respectively would be good theoretical bases for such microworlds. These microworlds should be integrated with each other where possible.

4.10 Research into interfaces based on group theory for other domains

An interesting theoretical (and perhaps practical) exploratory research project would be to investigate whether interfaces based on (ubiquitous) group theoretic characterisations of other domains may be designable for translating proximity relationships that happen to be hard to intuit in the abstract into haptic, visual, easily assimilable form.

4.11 Expanding the MC framework into a full guided discovery tutor

The MC framework for tutoring music composition relates the following components;

- an interface (Harmony Space)
- appropriate forms of knowledge at different levels (song, style and plan)
- ways of manipulating the knowledge (PLANC and Harmony Logo)

However, MC lacks the components needed to turn it into a full guided discovery

tutor. In terms of the 'traditional' Intelligent Tutoring Systems trinity, the components already defined correspond to roughly to the "domain expertise", but student models and teaching expertise are lacking. An interesting research project would be to expand MC using such traditional components as;

- a semantic network of musical concepts,
- user and curriculum models,
- stereotypical user models,
- appropriate teaching strategies and
- adaptive microworld exercises,

in order to implement *active* guided discovery learning teaching by the system. Suitable teaching strategies to investigate include the 'Paul Simon' method and 'most desirable path' method outlined in the third part of the thesis, as well as versions of some of the domain independent teaching strategies used in Dominie (Spensley and Elsom-Cook, 1988).

4.12 More ambitious variant designs of Harmony Space

An interesting large scale project would involve implementing and testing a number of variants of Harmony Space to explore various educational possibilities. A few of the most interesting variants are listed below.

- *Non-tonal versions.*

Versions of Harmony Space similar to the present design, but with 20, 30 and 42 fold divisions of the scale. These divisions are identified by Balzano (1980) as being of particular theoretical interest. Such a version of Harmony Space would allow student composers to experiment with, and assess the resources of, these non-tonal scales.

- *Pentatonic version*

A version of Harmony Space similar to the present design but with the ability to 'swap foreground and background' (Fig 2). This would make it possible to impose a restriction allowing only "black" notes to be played, as opposed to the current

optional restriction involving "white" notes. The aim of this version would be to allow student to assess harmonic resources of the pentatonic scale (e.g. tendency to perceive tonal centres) and compare them with those of the tonal system.

• *Just-tuned; non-repeating space version*

It is proposed that a version of Harmony Space be constructed based on Longuet-Higgins non-repeating space, connected via MIDI to a synthesiser capable of playing in just intonation (e.g. a Yamaha DX7 II), to help student compare different characterisations of pitch.

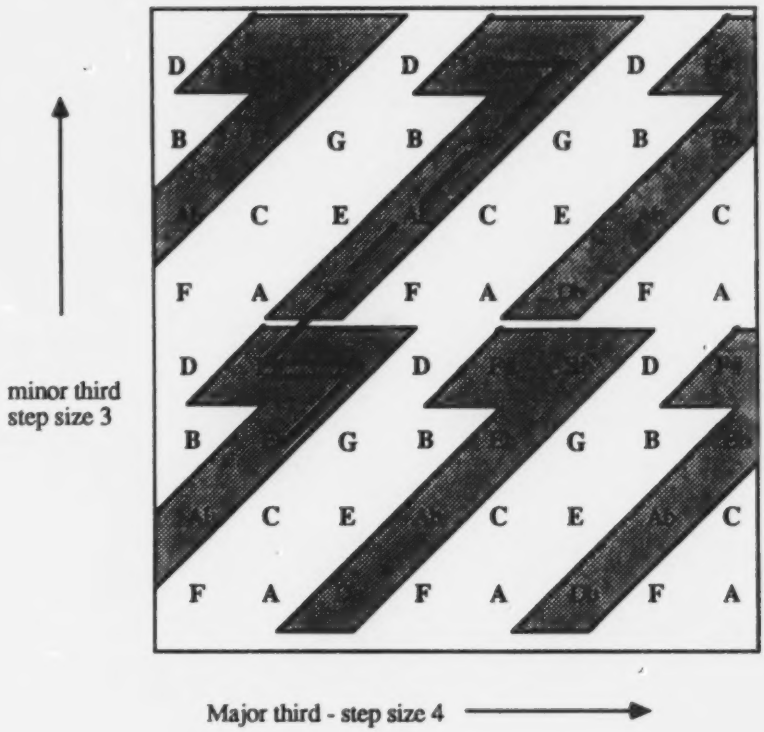


Fig 2 Assessing the harmonic resources of the pentatonic scale: four perfect fifth diads; one major and one minor triad; no privileged harmonic centre.

• *Performance event version*

A versions of Harmony Space similar to the present design, but with note location controlled by the location of people in a room or on a grid. Optionally, note names and key windows could be displayed on the floor using appropriate technology. The aim of this versions would be to allow children and other students to experience and

control chord progressions in a principled way using their entire bodies.

• *Multitrack version with graphic playback*

A versions of Harmony Space similar to the present design, but allowing chord sequences to be stored and subsequently played back synchronised with synchronised sound and graphics. This version should also allow the building up of multitrack recordings by permitting new tracks to be recorded while old ones are played back. Editing facilities would be particularly useful. Note that melodies, bass lines and accompaniments could be built up in this way simply by setting chord arity and pitch register appropriately on successive passes. In fact, given a multitrack or collaborative (see below) versions of Harmony Space, there is no reason why it should not be used for the playing, analysis and modification of polyphonic, as opposed to homophonic music.

• *Collaborative version*

A versions of Harmony Space similar to the present design, with multiple mice to allow several students to perform together, for example splitting responsibility for melody, bass line and accompaniment, or performing polyphonic lines. Responsibility could be split in various ways using multiple limbs, multiple users or multiple tracking.

• *Rubber band version*

A version of Harmony Space similar to the present design, but allowing chord sequences to be generated by preparing 'rubber-band' line-and-arrow diagrams to show the chord sequence 'where to go'. ('Rubberbanding' is a technique for drawing with a computer and pointing device that allows sequences of straight lines to be drawn quickly and accurately without demanding continuous physical precision of gesture. (For example, MacDraw is well-known drawing program that uses 'rubber-banding' techniques.) Such a version of Harmony Space could allow a sequence of chords to be drawn directly on the screen as a line diagram like fig 13, and edited quickly and easily prior to being played back. This contrasts with the real-time, immediately sounding methods of gestural control described so far. The

line drawings could be annotatable to record additional control information such as chord arity and rhythmic information. This could allow complex and fast chord sequences to be played without a need for great dexterity. It would also be valuable if chord sequences not necessarily prepared in this way could be modified and edited using this technique. This could greatly assist in the modification of existing pieces as an educational technique.

• *Version with rhythmic controllers*

A version of Harmony Space similar to the present design, but allowing the performance to be modulated by rhythmic figures selectable from an on-screen palette. A useful addition would be an drum interface to allow live, complex control of rhythm by a second player.

4.13 Interface design techniques to promote accessibility

Harmony Space exemplifies and combines an unusually high density of generally applicable interface techniques for making abstract or inaccessible entities and relationships accessible. Theories of interface design are scarce, but it is suggested that Harmony Space provides a useful precedent for future research on interfaces designed to promote accessibility. Four techniques simultaneously exemplified by Harmony space can be identified as follows;

• *Make visible a representation underlying a theory of the domain.* In the case of Harmony Space, C_{12} (and the structures implicit in it) are made concretely visible and directly manipulable. In TPM (Eisenstadt and Brayshaw, 1988), the proof trees underlying Prolog execution are made visible (but not directly manipulable).

• *Map a task analogically from a modality in which a task is difficult or impossible for some class of user into another modality where it is easy or at least possible to perform.* Apart from Harmony Space, two other interfaces that use this technique are Soundtrack, (Edwards, 1986) and TPM (Eisenstadt and Brayshaw, 1988). Soundtrack maps a WIMP graphic interface into the auditory modality so that blind and partially sighted users can use it. TPM (The Transparent Prolog Machine) dynamically maps Prolog

proof trees into graphical trees, to allow relative novices to debug programs and understand program execution more easily.

• *Use a single, uniform, principled metaphor to render abstract, theoretical relationships into a form that can be concretely experienced and experimented with.* In the case of Harmony Space, the abstractions are those of music theory. A striking recent example of another interface using this technique is ARK (the Alternative Reality Kit) (Smith, 1987). In ARK, *all* objects represented in the interface have position, velocity and mass and can be manipulated and 'thrown' with a mouse-driven 'hand'. The laws of motion and gravity are enforced but may be modified, allowing students to perform alternative universe experiments. Much of the experiences it deals with cannot normally otherwise be experienced in the original modality (due to travel restrictions, friction and the fixity of universal constants) and must normally be approached using abstract formulae. ARK uses the abstractions of physics to create a world in which such local barriers can be surmounted, allowing students autonomous access to purposive experiences.

• *Search for and exploit a single metaphor that allows principles of consistency, simplicity, reduction of short term memory load and exploitation of existing knowledge to be used.* The aim is to make a task normally difficult for novices easy to to perform. Harmony Space does this (as explained in detail above), but perhaps the best known example of such an interface is the Xerox Star (and latterly Macintosh) interfaces, which exploit a shared office metaphor and uniform commands consistently across a range of contexts.

We will summarise the distribution of these techniques[2] in the collection of interfaces noted above, making use of a table. We will use the following abbreviations;

TPM the Transparent Prolog Machine (Eisenstadt and Brayshaw, 1988),
ARK the Alternative Reality Kit (Smith, 1987),

[2] There are other interface design techniques that promote accessibility, including some without generally agreed names.

ST Soundtrack (Edwards, 1986),

HS Harmony Space (Holland, 1988)

In some cases, a technique is given as not used where it is used only to a slight or doubtful degree.

	HS	TPM	ARK	ST	XS
Theory-based representation visualisation	Y	Y			
Single metaphor direct manipulation	Y		Y	Y	Y
Cross-modality analogical mapping	Y			Y	
Exploiting existing knowledge	Y		Y		Y

An interesting research project would be to analyse these and other highly empowering interfaces, with the aim of explicitly characterising and generalising the techniques they use, and establishing the extent to which they can be applied to wider domains.

We have presented a number of ways of using artificial intelligence to foster and encourage music composition by beginners, particularly those without access to formal instruction. We have argued that the research has implications for artificial intelligence and education in general, because the support of open-ended approaches is relevant to any complex domain. We hope that the present work will contribute to the new research project that we have just outlined: whose ultimate aim would be to devise a theoretical framework for the design of highly empowering interfaces.

References

Aldwell, E. and Schacter, C. (1978) Harmony and Voice Leading. Harcourt Brace Jovanovich, Inc., New York,

Ames, C. and Domino, M. (1988) A Cybernetic Composer Overview. Proceedings of the First International Workshop on Artificial Intelligence and Music, AAAI 1988, St. Paul, Minnesota.

Baily, J. (1985) Music Structure and Human Movement. In Howell, P., Cross, I. and West, R. (1985). (Eds.) *Musical Structure and Cognition,* Academic Press, London.

Baeker, R.M and Buxton, W.A.S. (1987) Readings in Human Computer Interaction Morgan Kaufman, Los Altos.

Baker, M.J. (1988) An Architecture of an Intelligent Tutoring System for Musical Structure and Interpretation. Proceedings of the First International Workshop on Artificial Intelligence and Music, AAAI 1988, St. Paul, Minnesota.

Baker, M.J.(in press) Arguing with the tutor: a model for tutorial dialogue in uncertain knowledge domains. In Guided Discovery Learning Systems, Ed. Elsom-Cook, M. Chapman and Hall, (in press).

Balzano, G. J. (1980) The Group-theoretic Description of 12-fold and Microtonal Pitch systems. In *Computer Music Journal* Vol. 4 No. 4. Winter 1980.

Balzano,G.J. (1982) The pitch set as a level of description for studying musical pitch perception. In Clynes, M. (Ed.) Music , Mind and Brain: the neuropsychology of music. Plenum, New York.

Balzano, G.J. (1987) Restructuring the Curriculum for design: Music, Mathematics and Psychology. In *Machine Mediated Learning,* **2**: 1&2.

Bamberger, Jeanne (1972) Developing a new musical ear: a new experiment. Logo Memo No. 6 (AI Memo No. 264), A.I. Laboratory, Massachusetts Institute of Technology.

Bamberger, Jeanne, (1974) Progress Report: Logo Music Project. A.I. Laboratory,

Massachusetts Institute of Technology.

Bamberger, Jeanne, (1975) The development of musical intelligence I: Strategies for representing simple rhythms. Logo Memo 19, Epistemology and Learning Group, Massachusetts Institute of Technology.

Bamberger, Jeanne (1986) Music Logo. Terrapin, Inc. Cambs. Mass.

Barr, A. and Feigenbaum, E., Eds., (1982) The Handbook of Artificial Intelligence, Vol. II. Pitman, London.

Beckwith, R.S. (1975) The Interactive Music Project at York University. Technical report, Ontario Ministry of Education, Toronto.

Behan, Seamus (1986) Member of rock group "Madness". Quoted in 'Keyboard practicing hints', 'Making Music', issue 1, April 1986. Track Record publishing Ltd, St Albans.

Björnberg, Alf (1985) On Aeolian Harmony in Contemporary popular music. Unpublished manuscript.

Bloch, G. and Farrell, R. (1988) Promoting creativity through argumentation. Proceedings of the Intelligent tutoring systems Conference, Montreal.

Blombach, A. (1981) An Introductory Course in Computer-Assisted Music Analysis: The Computer and the Bach Chorales. In Journal of Computer-based instruction, Vol 7 No 3 February 1981.

Brayshaw, M. and Eisenstadt, M. (1988) Adding Data and Procedural Abstraction to the Transparent Prolog Machine. HCRL Technical Report 31, Open University. Submitted to 5th Logic Programming Conference, Seattle, Washington 1988.

Burton R.R. and Brown J.S.,1982. An Investigation of Computer coaching for informal learning activities. In Intelligent Tutoring Systems, Sleeman D. and Brown J.S., Eds. Academic Press 1982.

Buxton, W., Sniderman, R., Reeves, W., Patel, S., and Baecker, R. (1979). The evolution of the SSSP score editing tools. Computer Music Journal 4, 1, 8-21.

Buxton, W. Patel, S., Reeves, W. and Baeker, R. (1981). Scope in Interactive Score Editors. Computer Music Journal, Vol. 5 No. 3 p.50-56

Buxton, W. (1987) The Audio Channel. In Readings in Human Computer

Interaction, Baeker, R.M. and Buxton W.A.S., Eds. Kaufman, Los Altos.

Buxton, W. (1988) The "Natural " Language of Interaction : A perspective on non-verbal dialogues. Invited talk, Nato Advanced Study Workshop, Cranfield, England.

Buxton, W. (1987) "User interfaces and Computer Music: the State-of-the-Science of the Art". Seminar given at the Open University, Thursday 24th September in the Audio Visual Room.

Camilleri, L. (1987) Letture Critiche. In Bequadro, numero 26, anno, aprile-guigno 1987.

Campbell, Laura. (1985) Sketching at the Keyboard. Stainer and Bell, London.

Cannel, W. and Marx, F. (1976) How to play the piano despite years of lessons. What music is and how to make it at home. Doubleday and Company, New York.

Carbonell, J.R. (1970) Mixed-initiative man-computer instructional dialogues. BBN Report No. 1971, Bolt Beranek and Newman, Inc., Cambridge, Mass.

Carbonell, J.R. (1970b) AI in CAI: an artificial intelligence approach to computer-aided instruction. IEEE Transactions on Man-Machine systems. MMS 11 (4): 190-202.

Cassirer, E. (1944) The concept of group and the theory of perception, Philos. Phenomenolog. Res., 5:1-36.

Chadabe, J. (1983) Interactive Composing. In Proceedings of the 1983 International Computer Music Conference. Computer Music Association, San Francisco.

Citron, S. (1985) Song-writing. Hodder and Stoughton, London

Clancey, W.J. (1983) The advantages of abstract control knowledge in expert system design. Proceedings of the national conference on expert system design, Washington D.C., pages 74-78.

Clancey, W.J. and Letsinger, R. (1981) Neomycin: reconfiguring a rule-based expert system for application to teaching. Proceedings of the Seventh International Joint Conference on Artificial Intelligence, Vancouver. pages 829-835.

Clancey, W. J. (1987) Knowledge-Based Tutoring. The GUIDON Program. The MIT Press.

Coker, J. (1964) Improvising Jazz. Prentice Hall, New Jersey.

Cooper, Paul (1981) Perspectives in Music theory. Harper and Row, London.

Cope, D. (1988) Music: The Universal Language. Proceedings of the First International Workshop on Artificial Intelligence and Music, AAAI 1988, St. Paul, Minnesota.

Cork, C. (1988) Harmony by Lego bricks. Tadley Ewing Publications, Leicester.

Cross, I., Howell, P. and West, R. (1988) Linear Salience in the perception of counterpoint. Presentation given at the First Symposium on Music and the Cognitive Sciences, IRCAM, Paris.

Desain, Peter and Honing, Henkjan (1986) LOCO: Composition Microworlds in Logo. Proceedings of International Computer Music Conference.

Desain, P. (1986) Graphical programming in Computer Music: A Proposal. In Proceedings of International Computer Music Conference.

Deutch, D. (1982) The Cognitive Psychology of Music. Academic Press, New York.

Dowling, W.J. (1986) Music Cognition. Academic Press, New York.

DuBoulay, B. (1988) Towards more versatile tutors for programming. Invited talk, Nato Advanced Study Workshop, Cranfield, England.

Ebcioglu, Kemal (1986) An expert System for Harmonising Four-part Chorales. Proceedings of International Computer Music Conference.

Edwards, A. D. N. (1986) Integrating synthetic speech with other auditory cues in graphical computer programs for blind users, Proceedings of the IEE International Conference on Speech Input and Output, London, (1986).

Edwards, A. D. N. and O'Shea T. (1986) Making graphics-based programming systems usable by blind people, Interactive Learning International, 2, 3, (1986).

Edworthy J. and Patterson, R. D.,(1985) Ergonomic factors in auditory systems. In Brown I. D. (Ed.), Proceedings of Ergonomics International 1985, Taylor and Frances, (1985).

Eisenstadt, M. and Brayshaw, M. (1988) The Transparent Prolog Machine, (TPM): an execution model and graphical debugger for logic programming. Journal of Logic Programming 1988 .

Elsom-Cook, M. (1984) *Design Considerations of an Intelligent Tutoring System for Programming Languages*. unpublished PhD, University of Warwick.

Elsom-Cook, M. (1986) Personal communication in a tutorial with Simon Holland, concerning 'People Mats'. Geoffrey Crowther Building, Open University, Milton Keynes.

Elsom-Cook, M. (1988) The ECAL authoring environment. Third International Symposium on Computer and Information Science, Ismir.

Elsom-Cook, M (in press). (Eds.) Guided Discovery Learning Systems. Chapman and Hall, London.

Eno, Brian, 1978. Oblique Strategies. Opal, London.

Friberg, A. and Sundberg, J. (1986), Rulle. In Proceedings of the 1986 International Computer Music Conference, Computer Music Association, San Francisco.

Fry, Chris (1984) Flavours Band: A language for specifying musical style. Computer Music Journal Vol. 8 No. 4.

Foss, C.L. (1987) Learning from errors in algebraland, Technical report IRL 87 003 Institute for research on learning, Palo Alto.

Forte, A. (1979) Tonal Harmony in concept and practice. Holt, Rinehart and Winston, New York.

Fuerzeig, W. (1986) Algebra Slaves and Agents in a Logo-based mathematics curriculum. In Instructional Science 14 (1986) 229-254

Fugere, B.J., Tremblay, R. and Geleyn, L. (1988) A model for developing a tutoring system in Music. Proceedings of the First International Workshop on Artificial Intelligence and Music, AAAI 1988, St. Paul, Minnesota.

Fux, J.J. (1725/1965) Gradus ad Parnassum. Annotated and translated by A.Mann and J.Edwards, Norton, New York. (First published in Vienna, 1925).

Gaver, W. W. (1986) Auditory icons: using sound in computer interfaces, Human-

Computer Interaction, 2, 2, pp 167-177, (1986).

Gilmour, D. and Self, J. (1988). The application of machine learning to intelligent tutoring systems, In Self, J. (ed.) Artificial Intelligence and Human Learning (1988), Chapman and Hall, London.

Goldstein, I.P. (1982) The Genetic Graph: a representation for the evolution of procedural knowledge. In Sleeman and Brown, (eds.) Intelligent Tutoring Systems. Academic Press.

Great Georges Project (1980) The Blackie Christmas Oratorio: The Birth of a building: Symphony In Sea

Greenberg, Gary (1986) Computers and Music Education: A Compositional Approach. Proceedings of International Computer Music Conference.

Gross, D. (1984) Computer applications to Music theory: a retrospective. Computer Music Journal 8,4, Winter 1984.

Colles, H.C. (1940) (Ed.) Groves Dictionary of Music and Musicians. Macmillan and Co., London.

Gutcheon, P. (1978) Improvising rock piano, Consolidated music publishers, New York

Harpe, W. (1976) (Ed.) 7•UP. Great Georges Community Cultural Project (The Blackie) Liverpool 1, England.

Hayes, C, and Sutherland, R. (1989) Logo Mathematics in the classroom. Routledge, London.

Hofstadter, Douglas R. (1979) Godel, Escher, Bach: an eternal golden braid. Penguin Books, London.

Hofstetter, F. (1975) Guido: an interactive computer-based system for improvement of instruction and research in ear-training. Journal of Computer-Based Instruction, Vol 1, No. 4, pages 100-106.

Hofstetter, F. (1981) Computer-based aural training: the Guido system. Journal of Computer-Based Instruction, Vol 7, No. 3, pages 84-92.

HM Inspectorate (1985). Music from 5-16 (Curriculum Matters 4). HMSO, London.

Holland, S. (1984) Unfolding explanations at appropriate levels of detail. Unpublished MSc Thesis. Department of Psychology, University of Warwick.

Holland (1984b) Can machines really learn new knowledge? Unpublished manuscript, Department of Psychology, University of Warwick.

Holland, S. (1985) "How computers are used in the teaching of music and speculations about how artificial intelligence could be applied to radically improve the learning of composition skills". OU IET CITE Report No. 6.

Holland, S. (1986a) Speculations about how Artificial Intelligence could be applied to contribute to the learning of music composition skills. In AISB Quarterly, Spring 1986.

Holland, S. (1986b) "Design considerations for a human computer interface using 12-tone three-dimensional Harmony space to aid novices to learn aspects of harmony and composition" OU IET CITE report no. 7.

Holland, S. (1987a) Direct Manipulation tools for novices based on new cognitive theories of harmony. pp 182 -189 Proceedings of 1987 International Computer Music Conference.

Holland, S. (1987b) Computers in Music education. In "Computers in Education 5-13" (1987) Eds. Jones, A. and Scrimshaw, P. Open University Press, Milton Keynes.

Holland, S. (1987c) "A knowledge-based tutor for music composition" OU IET CITE Report No. 16.

Holland (1987d) "A knowledge-based tutor for music composition". Talk presented at the Third International Conference on AI and Education, LRDC, Pittsburgh, 1987.

Holland, S. (1988) "New issues in applying two cognitive theories of harmony" Poster presentation at International Symposium on Music and the Cognitive Sciences IRCAM (Institute de Recherche et Coordination Acoustique/Musique) Paris 14-18 March 1988.

Holland, S. and Elsom-Cook, M. (in press) Architecture of a knowledge-based tutor for music composition. In Guided Discovery Learning Systems, Ed. Elsom-Cook, M. Chapman and Hall, (in press).

Honing, Henkjan (1989) Personal communication. Telephone call by Henkjan from Lewisham to the author at the Open University concerning a reference for Miller Puckette's Patcher, and comments on a draft of Part II.

Howell, P., Cross, I. and West, R. (Eds.), (1985). *Musical Structure and Cognition*. Academic Press, London.

Huron, David (1988) The perception of polyphony. Presentation given at the First Symposium on Music and the Cognitive Sciences, IRCAM, Paris.

Hutchins, E.L., Hollans, J.D. and Norman, D.A. (1986) "Direct Manipulation Interfaces" in Norman, D.A. and Draper, S. (Eds.), User Centred System Design: New perspectives on Human Computer Interaction, Erlbaum, Hillsdale, NJ.

Illich, Ivan. 1976. Deschooling society. Pelican, London.

Jairazbhoy, N. (1971) The ragas of North Indian music. Wesleyan University Press, Middletown.

Johnson-Laird, P. N. (1988) The computer and the mind, Fontana, London.

Johnson-Laird, P. N. (1989) Personal communication. Intervention from the audience during seminar entitled"Harmony Space", given by S.Holland, 31st January 1989, in the Cambridge University "Seminar on Music and Mind" series at the University Music School.

Jones, Kevin (1986) Real Time Stochastic Composition and Performance with Ample. Proceedings of International Computer Music Conference.

Jones, Kevin (1984) Exploring Music with the BBC micro and Electron. Pitman, London.

Jones, Barry (1989) Personal communication. Criticisms of draft of chapter 9 kindly made to the author by Dr. Jones in his office in the Music Department, Open University, third week of July 1989.

Kasner, E. and Newman, J. (1968) Mathematics and the Imagination. Pelican Books, London.

Kilmer, A.D., Crocker, R.L. and Brown, R.R. (1976) Sounds from silence: recent discoveries in ancient Near East music. Bit Enki Publications, Berkely.

Krumhansl, C.L. (1979) The psychological representation of musical pitch in a tonal

context. Cognitive Psychology, 1979, 11, 346-374.

Krumhansl, C.L. (1987) Formal properties of various scale systems. Second Conference on Science and Music, City University, London.

Kurland, D.M., Pea, R.D., Clement, C., and Mawby, R. (1986) A study of the development of programming ability and thinking skills in high school students. Journal of Educational Computing Research. Vol 2, No. 4 1986 pp429-458.

Lamb, Martin (1982) An interactive Graphical Modeling Game for teaching Musical Concepts. In Journal of Computer-Based Instruction Autum 1982.

Laske, O.E. (1978) Considering human memory in designing user interfaces for computer music In Computer Music Journal 2,4, 39-45.

Laske, O.E (1980) Towards an explicit Cognitive Theory of Musical Listening Computer Music Journal Vol 4,2, pp73-83

Laurillard D M, (1988) Evaluating the contribution of information technology to students' learning. In Proceedings of the Conference Evaluating IT in Arts and Humanities Courses, Birmingham, May 1988.

Lenat D.B. and Brown J.S. (1984) Why AM and Eurisko appear to work. In Artificial Intelligence 21 31-59.

Lenat, D.B. (19[illegible]2) EURISKO: A program that learns new heuristics and domain concepts. In Artificial Intelligence 21, pp 61-98.

Lerdahl, Fred and Jackendoff, Ray (1983) A Generative Theory of Tonal Music. MIT Press, London.

Lerdahl, F. and Potard, Y. (1985) A computer aid to composition. Unpublished Manuscript.

Levitt, D.A. (1981) A Melody Description System for for Jazz Improvisation. Unpublished Master's Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

Levitt, David (1984) Constraint Languages. In Computer Music Journal Vol. 8, No.1.

Levitt, D.A. (1985) A Representation for Musical Dialects. Unpublished PhD Thesis, Department of Electrical Engineering and Computer Science, Massachusetts

Institute of Technology.

Levitt, D. H and Joffe, E. (1986). *Harmony grid* .Computer software, MIT.

Longuet-Higgins, H.C. (1962) Letter to a Musical Friend. In Music Review, August 1962 pp 244-248.

Longuet-Higgins, H.C. (1962b) Second Letter to a Musical Friend. In Music Review, November 1962.

Longuet-Higgins and Lee (1984) The rhythmic interpretation of monophonic music. In Music Perception Vol 1 No. 4 424-441
(chapter 1 page 7)

Loy, G. (1985) Musicians make a standard: The MIDI phenomenon. Computer Music Journal 9(4), 8-26.

Loy and Abbott, (1985) Programming Languages for Computer Music Synthesis, Performance and Composition. In ACM Computing Surveys Volume 17, No. 2 June 1985, Special Issue on Computer Music.

McAdams, S. (1987) "Organisational Processes in Musical Listening. Talk at Second Conference on Science and Music, City University, September 19th 1987.

McAdams, S. and Bregman, A. (1985). Hearing Musical Streams. In 'The Foundations of Computer Music', Eds. Roads C. and Strawn, J. MIT Press, London.

Making Music (1986) Three chord trick. Issue 3 June 1986 (and associated series) Track record publishing Ltd. London

Mathews, M.V. with Abbott, C.A. (1980) The Sequential Drum. In Computer Music Journal Vol 4, No 4 Winter 1980.

Mehegan, J. (1959) *Jazz Improvisation I: Tonal and Rhythmic Principles*. Watson-Guptil Publications, New York.

Mehegan, J. (1960) *The Jazz Pianist: Book Two*. Sam Fox Publishing Company, London.

Mehegan, J. (1985) Improvising Jazz piano. Amsco publications, New York

Minsky, Marvin (1981) Music, Mind and Meaning. In Computer Music Journal

Vol.5 No.3 p.28 - 44

Minsky, Marvin (1981b) A framework for representing knowledge. In Haugeland, J. (Ed) Mind design, The MIT press, Cambridge Mass.

Morely, T. (1597) A plain and easy introduction to Practical Music. (Ed. Harman, R.A. (1952) Dent, London.

Moyse, R. (1989) Knowledge negotiation implies multiple viewpoints. Proceedings of 4th International Conference on AI and Education. IOS, Amsterdam.

Nelson, T. (1974) Computer Lib/Dream Machines. The Computer Bookshop, Birmingham.

Newcomb, Stephen R. (1985) LASSO: An intelligent Computer Based Tutorial in Sixteenth Century Counterpoint. In Computer Music Journal Vol. 9 No.4.

Orlarey, Yann (1986) MLOGO - a MIDI Composing Environment for the Apple IIe. In, Proceedings of International Computer Music Conference. Computer Music Association, San Francisico.

O'Shea, Tim and Self, John (1983). Learning and Teaching with Computers. Prentice-Hall, London.

O'Shea, T., Evertz, R., Hennessey, S., Floyd, A., Fox, M. and Elsom Cook, M. (1987) Design Choices for an intelligent arithmetic tutor. Published in Self, J. (Ed.) Intelligent Computer Assisted Instruction. Methuen/Chapman & Hall.

O'Malley, C.E. (in press) Graphical interaction and Guided Discovery Learning, Ed. Elsom-Cook, M. Chapman and Hall, (in press).

Papert, S. (1980) Mindstorms. The Harvester Press, Brighton.

Paynter, J. (1982) Music in the secondary School curriculum. Cambridge University Press, London.

Pierce, J.R. (1983) The Science of Musical Sound, Scientific American Books, New York.

Pope, S.T. (1986) The Development of an Intelligent Composer's Assistant Interactive Graphics Tools and Knowledge Representation for Music. Proceedings of International Computer Music Conference.

Pearson, G. and Weiser, M. (1986) Of moles and men: the design of foot controls for work-stations. In proceedings of the International Conference of Computers and Human Interaction (CHI) 1986. ACM/SIGCHI, Boston.

Pennycook, B.W. (1985a) Computer-Music Interfaces - a Survey. In ACM Computing Surveys, Vol. 17 No 2 June 1985.

Pennycook, B.W. (1985b) The role of graphics in computer-aided music instruction (part 1). In the Journal of the electroacoustic msuc association of Great Britain, Vol. 1 No 3 1985.

Pennycook, B.W. (1985c) The role of graphics in computer-aided music instruction (part 2). In the Journal of the electroacoustic msuc association of Great Britain, Vol. 1 No 4 1985.

Pratt, G. (1984). *The Dynamics of Harmony*. Open University Press, Milton Keynes.

Puckette, M. (1988) The Patcher. Proceedings of the International Computer Music conference 1988. Feedback papers 33, Feedback Studio Verlag, Cologne.

Pulfer, J.K. (1971) Man-machine interaction in creative applications. International Journal of Man-machine Studies, 3, 1-11.

Reck, David (1977) Music of the Whole Earth. Scribner's, New York.

Rich, E.A. (1979) User Modelling via stereotypes. Cognitive Science 3 pp 329-354.

Richer, M.H. and Clancey, W.J. (1985). GUIDON_WATCH: a graphic interface for viewing knowledge-based systems. IEEE Computer GRaphics and Applications, Vol. 5, no. 11, pp. 51-64

Roads, C. (1981) Report on the IRCAM Conference: The Composer and the Computer. In Computer Music Journal Vol. 5 No. 3.

Roads, C. (1984) An overview of music representations. In Olschki, L.S., (Ed.) Proceedings of the Conference on Music grammars and Computer Analysis. Quaderni della Rivista Italiana di Musicilogia, Firenze.

Roads, C. (1985) Research in Music and Artificial Intelligence. In ACM Computing Surveys, Vol. 17 No 2 June 85.

Roads, C. and Strawn, J. (1985) (Eds) The Foundations of Computer Music. MIT Press, Cambridge Mass.

Rodet, X. and Cointe,P. (1984) Formes: Composition and Scheduling of Processes. In Computer Music Journal 8,3 pp 32-50.

Rosen, C. (1971) The Classical Style. Faber, London.

Rutter. J. (1977) Principles of form. (The Elements of Music, Unit 13). The Open University Press, Milton Keynes.

Sanchez, M., Joseph, A., Dannenberg, R., Miller,P., Capell,P. and Joseph,R. (1987). Piano tutor: An Intelligent Keyboard Instruction System. Proceedings of the First International Workshop on Artificial Intelligence and Music, AAAI 1988, St. Paul, Minnesota.

Schoenberg, A. (1954) Structural functions of Harmony. Ernest Benn Ltd, London.

Schoenberg, A. (1967) Fundamentals of music composition. Faber and Faber, London.

Scholz, Carter (1989) Notation software ahead: proceed with caution. In Keyboard, April 1989.

Self, J. A. (1974) Student models in computer-aided instruction, International Journal of Man-machine studies, Vol. 6, pp. 261 - 276.

Self, J. (1987) (Ed.) Intelligent Computer Assisted Instruction. Methuen/Chapman & Hall, London.

Self, J. A., (1988) The use of belief systems for student modelling, Internal CERCLE report, University of Lancaster

Self, J.A. (1988a) (Ed) Artificial Intelligence and Human Learning. Intelligent Computer-Aided Instruction. Chapman and Hall, London.

Sharples, Mike (1983) The use of Computers to Aid the Teaching of Creative Writing. In AEDS Journal.

Shepard, R.N. (1982) Structural representation of musical pitch. In The psychology of music (D. Deutch, Ed.) Academic press, New York.

Shneiderman, B. (1982). The future of interactive systems and the emergence of

direct manipulation. *Behaviour and Information Technology* **1** 237-256.

Simon, H. and Sumner, R. (1968) Pattern in music. In Formal Representations of Human Judgement. B. Kleinmutz, Ed Wiley, New York.

Sleeman D. and Brown J.S.(1982) (Eds) Intelligent Tutoring Systems. Academic Press 1982.

Sloane, B., Levitt, D. Travers, M. So, B. and Cavero, I (1986). Hookup 0.80. Prototype undistributed computer software for the Apple Macintosh. Media Laboratory, MIT, Cambridge, Mass.

Sloboda, J. (1985). *The Musical Mind: The Cognitive Psychology of Music.* Clarendon Press, Oxford.

Smith, D.C., Irby,C., Kimball, R., Verplank, W., and Harslem, E. (1982) Designing the Star User Interface. Byte 7(4), April 1982 242-282.

Smith, R.B. (1987) "Experiences with the Alternate Reality Kit: An example of the tension between literalism and Magic". Proceedings of the Conference on Human Factors in Computer Systems, Toronto, April, 1987.

Smoliar, S. (1980) A computer aid for Schenkerian analysis. In Computer Music Journal 4,2.

Sorisio, L. (1987) Designing an Intelligent Tutoring System in Harmony. Proceedings of International Computer Music Conference, Computer Music Association, San Francisco.

Spensley, F. and Elsom-Cook, M. (1988). Dominie: Teaching and assessment strategies. Open University CAL Research Group Technical Report No. 74, Milton Keynes, Great Britain.

Spiegel, L. (1981) Manipulations of musical patterns In Proceedings of the 1981 Symposium on small computers in the arts. IEEE.

Spiegel, L. (1986) Music Mouse (Computer software for Apple Macintosh). Opcode systems.

Steedman, M. (1972) *The formal description of musical perception* . Unpublished Phd, University of Edinburgh.

Steedman, M. (1984) A generative grammar for jazz chord sequences. *Music*

Perception **2**:1.

Stevens A., Collins A. (1977) The Goal Structure of a Socratic Tutor, in Proceedings of the ACM Annual Conference.

Stevens A., Collins A., Goldin S.E. (1982) Misconceptions in Student's Understanding , in D.H.Sleeman, J.S.Brown (eds) Intelligent tutoring systems, pp. 13-24, Academic Press.

Sudnow, D. (1978) "Ways of the hand: the organisation of improvised conduct." Routledge and Kegan Paul, London.

Sussman, G.J. and Steele G.L. (1980) Constraints: A language for expressing almost hierarchical descriptions . AI Journal, Vol 14 (1980), p1-39

Tanner, P. (1971) Some programs for the computer generation of polyphonic music. Report ERB-862, National Research Council, Ottowa, Canada.

Tobias, J. (1988) Knowledge Representation in the Harmony Intelligent Tutoring System. Proceedings of the First International Workshop on Artificial Intelligence and Music, AAAI 1988, St. Paul, Minnesota.

Todd, N. (1985) A model of expressive timing in tonal music. Music Perception, 3:1

Thomas, Marilyn Taft (1985) VIVACE: A rule-based AI system for composition. In Proceedings of the 1985 International Computer Music Conference, Computer Music Association, San Francisco.

Truax, B. (1977) The POD system on interactive composition programs. Computer Music Journal 1,3, 30-39.

Twidale, M.B.(1988) Explicit planning and instantiation as a means to facilitating student-computer dialogue. Proceedings of the European Congress on Artificial Education and Training, Lille.

Twidale, M.B. (1989) Intermediate representations for student error diagnosis and support. Proceedings of 4th International Conference on AI and Education . IOS, Amsterdam.

Vercoe, B. (1975) Man-Machine interaction in creative applications. Experimental Music Studio, Massachussetts Institute of Technology, Cambridge, Mass.

Watkins, J. and Dyson, Mary C. (1985) On the perception of tone sequences and melodies. In Howell, Peter, Cross, Ian, and West, Robert (Eds.) Musical Structure and Cognition. Academic Press, London.

Weber, G. (1817) (Trans. 1851) Theory of Musical Composition London: Cocks.

Wenger, E. (1987) Artificial Intelligence and Tutoring Systems. Morgan Kaufmann Publishers Inc., Los Altos, California.

Wilding-White, R. (1961) "Tonality and Scale theory". Journal of Music Theory 5:275-286.

Winograd, T. (1968) Linguistics and the Computer Analysis of Harmony. Journal of Music Theory, 12. Pages 2-49.

Winsor, P. (1987) Computer assisted music composition. Petrocelli Books, Princetown.

Winston, P. (1984) Artificial Intelligence. Addison-Wesley, London.

Wittgenstein, L.(1961/1922) Tractatus Logico-Philosophicus. Routledge and Kegan Paul, London.

Wright, J.K. (1986) Auditory object perception: counterpoint in a new context. Unpublished MA Thesis, Faculty of Music, McGill University.

Xenakis, I. (1971). Formalised Music.Indiana University Press, Bloomington.

Xenakis, I. (1985) Music Composition Treks. In Roads, C. (1985) Composers and the Computer. Kaufmann, Los Altos.

Zicarelli, D. (1987). M and Jam Factory. In Computer Music Journal, Vol. 11, No. 4, Winter 1987

Zimmerman, T.G., Lanier, J., Blanchard, C., Bryson, S. and Harvill, Y. (1987) A Hand Gesture Interface Device. In proceedings of the Conference on Computer Human Interaction, ACM.

Musical references

Bach, J, S.(1968/~1722) Bach's prelude No 1 in C Major from the 48 Preludes and Fugues. Kalmus, New York. (Written around 1722.)

Berry, Chuck (1958) "Johnny B. Goode". Jewel Music, London (Chappell). Chess 1691.

Bowie, David (1983) "Let's Dance". EMI Music, London.

Clapton, Eric (1970) "Layla". Throat Music, London.

Collins, Phil (1981) "In the Air Tonight". Hit and Run Music, London.

Collins, Phil (1984) "Easy Lover", (Phil Collins, Phil Bailey and Mason East) MCA, Warner-Chappell, Hit and Run, London. First released 1985.

Coltrane, John (1959) "Giant Steps". Atlantic 1311.

Crawford, Randy (1978) "Street Life". Performed by Randy Crawford and The Crusaders, MCA 513. Written by Joe Sample and Will Jennings. Published by Rondor and Warner-Chappell, London.

Davis, Miles (1959) "So What" on LP "Kinda Blue". Columbia, CL 1355

Dylan, Bob (1968) "All along the Watchtower". Feldman, EMI, London.

Dire Straits (1978) "Sultans of Swing". Straight-Jacket/ Rondor, London

Gershwin, George and Ira (1930) "I Got Rhythm". New World Music Corporation, New York.

Lennon McCartney (1968) Birthday. The White Album. BIEM 8E 164 04 174A

Jobim, C.A. and Moraes, V. (1963) "Girl from Ipanema". MCA Music, London. On Getz, S. and Gilberto, A. (1963) Verve 68545.

Hartley, Keef (1969) "Outlaw" on LP "Half-Breed". Deram SML 1037.

Hendrix, Jimi (1980 re-issue) "Hey Joe" (popular arr. Jimi Hendrix). Polydor 2608001.

Hawkins, F., Holland, S., and Clarke, C. (1980) "Fox". © Ex post Facto. 8, Second Avenue, Liverpool 9.

Jackson, Michael (1982) "Billie Jean", on "Thriller". CBS Inc/Epic.

Kern-Hammerstein "All The Things You are". Chappell & Co. and T.B. Harms.

Miller, Steve (1984)"Abracadabra", Warner-Chappell (music publishers), London.

Soft Machine (1970) "Out-Bloody-Rageous" (Ratledge, M) on LP Soft Machine "Third". CBS 66246.

Police (1979) "Message in a Bottle" (Sting). Virgin Music. Issued on LP "Regatta de Blanc".

Rogers, R. and Hart, L.(~1934) "The lady is a tramp". Chappell, London. First Frank Sinatra recording, 1956.

Steely Dan (1974) "Pretzel Logic" on LP "Pretzel Logic". ABC, ABCL 5045

Taj Mahal (~1968) "Six Days on The Road" on "Fill Your Head with Rock" (various artists) CBS records.

Youmans, Harbach, Caesar (1924) "Tea for Two". Harms, Inc.

Walsh (1976) "I don't need another girl like you" © John Walsh, 54 Bradbury Road, Solihull.

Wonder, Stevie (1973) "Living for the City" on LP "Inner Visions". Tamla Motown STMA 8011.

Wonder, Stevie (1976) "Isn't she lovely" Jobete Music Co. Inc. and Black Bull Music Inc. On LP 'Songs in the Key of Life', Motown STML 60022.

Appendix 1: Harmony Space prototype implementation

This appendix describes the prototype implementation of Harmony Space.

Apart from the limitations and bugs discussed, the prototype satisfies the core abstract design of Harmony Space given in Chapter 5, in the Balzano configuration. But the abstract design is only concerned with design decisions at the level of abstraction of chapter 5 section 3.1. To build a particular implementation that satisfies the abstract design, decisions must be taken about such matters as the layout of menus used to control secondary functions and what controller is used to control particular pointing functions.

Since the prototype was originally intended to demonstrate the expressivity of the core abstract design, and not to carry out investigations with novices, the *secondary functions* in the prototype were implemented in a relatively easy-to-program fashion, regardless of their user-hostile, non-DM style.

We begin with an illustrated overview of the appearance and method of use of the prototype[1]. We then quickly summarise technical details of the implementation (hardware employed, etcetera). Following this, we look at the limitations and scope for improvement of the interface. Finally, we give a summary list of all of the commands supported. There are numerous illustrations and diagrams referred to, to be found at the end of this appendix.

1 General appearance and operation of prototype.

We start with a quick general overview of the appearance and method of use of

[1] Note - the material in this appendix inevitably overlaps to some extent material already seen in Chapter 5 and appendix 2. Identical or almost identical wording is used in parts, although the description is carefully tailored here to refer to the prototype implementation as opposed to the notional or minimal version.

the Harmony Space prototype. This overview draws on the account of Harmony Space already presented in Chapter 5, but with changes to tailor the description to the prototype exactly as implemented. Note that the prototype employs the Balzano theory of Chapter 7 rather than the Longuet-Higgins' theory of Chapter 5. This makes a difference to the configuration of notes on the screen and the shape of the key-windows.

A screen dump of the Harmony Space window in the Harmony Space prototype is shown in Fig 1. The Harmony Space window consists of a two-dimensional grid of circles, each circle representing a note. The notes are arranged in ascending major thirds (semitone steps of size 4) on the x-axis, and in ascending minor thirds (semitone steps of size 3) on the y-axis. As described in chapters 5 and 7, the notes of the diatonic scale 'clump' into box shaped regions on the screen. 'Key windows' are used to mark out these diatonic areas on the screen in white (Fig 1) [2]. The key windows can be slid as a group over the fixed grid of circles by means of the vertical, horizontal and diagonal arrow keys. (Figs 6,7 and 8 are screen dumps showing three different positions of the key window corresponding to the keys of the tonic, mediant and dominant of the starting key respectively.) A mouse is used to move the cursor around in the window. If the mouse button is pressed while the cursor is over a note circle, the circle lights up and the corresponding note sounds audibly. Letting go of the button normally causes the circle to go dark again, and is like letting go of a synthesiser key - the sound stops or decays appropriately (depending on the synthesiser voice currently selected [3] and how it responds to a midi note-off command). Similarly, if the button is *kept* depressed and the mouse is swept around, the appropriate succession of new notes is sounded and appropriately completed as the mouse enters and moves the note circles (but see the section on limitations, below).

[2] This is how the key areas are shaded in the prototype. Other things being equal, it appears to be a good rule to shade the key windows white and the chromatic areas dark, so as to echo the arrangement of black and white keys on a piano. This might marginally help skill transfer from Harmony Space to keyboards, given a keyboard display as in appendix 20. For technical reasons to do with the the difficulty of printing large expanses of black, we have presented the screen dumps here inverted black-for white. Since the actual screen dumps are over 1Mb each in size - too large to include in the thesis document - they have been redrawn manually.

[3] This is done externally at the synthesiser, though it would be trivial to arrange for patch selection to be done in some other way.

More generally, the mouse can be set to control the location of the root of a diad, triad, seventh or ninth chord. (We have called the number of notes in a chord the chord *'arity'*, borrowing from mathematical usage.) For any given chord arity, depressing the mouse button on a root causes all of the notes of the appropriate chord to light up and sound. (Fig 1 shows a I major triad being sounded in this way on the prototype.)

Apart from the mouse and arrow keys, other functions in the prototype are controlled by a set of function keys. (See figs 2 and 3 for diagrams showing the functions of the various control keys. Fig 2 is an overview of the layout, while Fig 3 shows the function of each key. There is also a summary list of key functions at the end of this appendix.) For reasons already indicated, the key-function mappings in the prototype are more or less arbitrary, cushioned by token attempts to make mnemonically helpful mappings. The prototype allows the user to select (by pressing the appropriate buttons - see figs 2 and 3) whether she wants (until further notice) chords of arity one, two, three, four, five, etcetera. (i.e., single notes, diads, triads, sevenths, ninths). In normal use, the prototype chooses the quality of each chord automatically in accordance with the prevailing chord-arity and the current location of the root relative to the current position of the key window. As different chord roots are triggered, the qualities of the chords vary appropriately according to the position of the root within the scale. So, for example, the chord on the tonic is normally a major triad (or major seventh if we have a prevailing chord-arity of sevenths) and the chord on the supertonic is normally a minor triad (or minor seventh if the arity is sevenths). We refer to chord qualities assigned in this way as the *default* chord qualities for the relevant arity and degree of the scale. This automatic assignment of quality to chords can be *overridden* in the prototype at any time by selecting the desired quality using the function keys (see figs 2 and 3). Such a manually selected quality affects only a single chord event - the very next one. After that, the assignments revert to the prevailing defaults, unless overidden again.

As we noted in Chapter 5, although the chord quality is normally chosen automatically in this way, the constraining effects of the walls of the key window provide a visible 'explanation' for the way in which the qualities are chosen. In

particular, the traditional rule for constructing 'scale-tone' chord qualities for any chord-arity on any degree of the diatonic scale (namely the 'stacking thirds' rule) has an exact visual analog on the screen of the prototype. In visual terms, the next note in the chord must always be chosen in the *outward-going* sense along whichever axis (north or east) the key window permits at that point. (The diatonic scale is such that there is always only one choice.) Fig 4 is a screen dump from the prototype that illustrates just this. It shows a chord being built up from a single note to a diad, third, seventh and ninth chord on the same root. (Normally, a chord is erased from the screen automatically when the triggering mouse button is released, but we have used an option (see figs 2 and 3) to leave a trace on the screen.) The two pointing devices (mouse and arrow keys for chords and key window respectively) in conjunction with the visible constraint metaphor make the prototype a powerful tool for the physical control of chord sequences by novices. With the two pointing devices the user can, for example, repeatedly play a chord root while moving the key window, and the chord quality will change appropriately. (Exactly this point is illustrated by figs 6,7 and 8. As the key window is moved from tonic to mediant to dominant, successive soundings of a chord of arity seventh on G are shown changing quality automatically from major seventh to minor seventh to half-diminished seventh.)

Alternative mappings of degree of scale to chord quality are also implemented in the prototype. For example, there are schemes implemented for dealing with the harmonic minor mode and dorian mode. Such alternative chord quality schemes can be selected until further notice simply by pressing the appropriate quality scheme key (see figs 2 and 3). Fig 5 shows a V chord being played in a minor key with the chord quality scheme for the harmonic minor in operation. This causes the V chord to be played as a cadential dominant - with the sharpened third of the modified chord visibly sticking outside the bounds of the key window. Apart from the automatic selection or momentary overiding of chord qualities, the other way of dealing with chord quality supported by the prototype is to fix chord quality until further notice (see Figs 2 and 3). We now complete this general overview by quickly noting the main other features of interest of the prototype, which are as follows. Chords are played in their root position. All of the *roots* are mapped into the compass of a single octave, with the rest of the chord in close

position on top of the root. The octave into which the roots are mapped can be varied by means of function keys (see figs 2 and 3). Roots outside of the key window (chromatic roots) can be muted or associated with any fixed arbitrary chord quality until further notice (see figs 2 and 3). The display can be adjusted either to show all notes of every chord, or the roots only. The display can also be adjusted so that the trail of roots that have been previously played are automatically erased as the chords are released (this is the norm), or left on the screen to show the history of the progression. (Fig 9 shows a screen dump from the prototype resulting from playing the chord sequence of 'Street Life' (Crawford, 1978) (see chapter 8 for the chord sequence) with the 'show root only' and 'leave trace' roots both engaged: the extended cycle of fifths progression is clearly visible. This completes the overview of the implemented prototype of Harmony Space. For diagrams of the command keys, see figs 2 and 3. For a listed summary of the key functions, see the end of this appendix.

2 Technical details of implementation

The prototype is implemented in Lucid Common Lisp and its object-oriented extension 'Flavours' using the DOMAIN Graphics Primitives Routines package (GPR) on an Apollo Domain DN300 workstation controlling a Yamaha TX816 synthesizer via a Hinton MIDIC RS232-to-MIDI converter. A simple midi driver for the Apollo was written in Pascal, and numerous low-level graphics routines were written in Lisp.

3 Limitations of implementation

The major limitations of the implementation are as follows. All of these are trivial, easily fixable problems from a programming point of view. The reason for documenting them rather than fixing them is that time constraints rule out the mechanical but laborious programming required to address them. Some of the bugs may seem too trivial to document, but they are recorded because they are far from trivial from the user's point of view. Some of them drastically affect the ease with which novices can currently use the prototype (see chapter 9). (To see an example of what a version of the interface might look like with some of these trivial but irritating problems addressed, see appendix 2.)

Non-existent note labelling. The note circles are not labelled at all. Ideally it should be possible to dynamically select from a choice of labellings for the note circles. Three obvious options would be alphabetic note names, degrees of the scale or semitone numbers. (See appendix 2.)

Some states and possible choices of state are not visible. There are no menus on screen to show possible choices of arity, chord quality and so forth. Similarly, there is no display (menu or otherwise) to show which of these states is currently selected. (See appendix 2 for a possible system of menus.)

Registration and misstriggering. The note circles as drawn, compared with as detected, are not quite in alignment. Occasionally the wrong chord triggers. It should be possible to play a sequence of notes in a single gesture by playing one note, keeping the cursor depressed, and playing any sequence of notes by threading through the spaces between the note circles and across chosen note circles. As the prototype stands, this can be done with a series of mouse button depressions, but not in a single, fluid gesture. The sensing radius of the note circles seem to be too large, and when sweeping out sequences with the button fixed down, the mouse continually bumps into unintended note circles. A related bug is when chords with chromatic roots have been muted and the mouse is swept with the button held down, for some reason the chromatic roots sound (this does not occur if each root is released before trying to play the next one).

Speed of response. It appears to take around 250 milliseconds for Harmony Space to respond to a mouse button depression. The prototype pauses once every few hundred events or so for half a minute for garbage collection. The fastest that the prototype can retrigger a chord is about once every half-second. If the button is held down and swept around, chords are produced slightly faster than one per second. Subjectively, the cursor seems to move in a slightly treacle-like fashion, and does not always seem to move exactly where it is directed. A response time of around 25 to 50 times faster would be adequate, since for many musical purposes, events separated by 10 milliseconds are perceived as instantaneous. Now that an exploratory prototype has been built, it would not be hard to address the speed problem by this factor simply by recoding more efficiently - for

example by declaring arrays and vectors and avoiding unnecessary message passing that was useful for the design and debugging phase.

Faults in trace mechanism. As we have already noted, it is possible in the prototype to have the whole chord displayed or the roots displayed only. It is similarly possible to have each chord erased immediately after it is played or left as a permanent trace. However there are one or two related shortcomings in these features as implemented. Firstly, due to a minor bug, it does not seem to be possible to get a permanent trace when the chord arity is one. Secondly, when lit-up note-circle traces are left, they are not really left permanently - they are simply not erased. Unfortunately, this means that if a note that has already been sounded is sounded again, then the trace becomes erased at that point.[4]

Lack of multiple pointing devices The use of one mouse to control root position, and function keys to control all other functions is very limiting. For a better notional arrangement of controllers, see appendix 2.

Lack of control of inversion, spacing and doubling. All chords are currently sounded in a close position. This can sound a bit wearying. See appendix 2 for some simple possible refinements.

Lack of control of position of octave break. All of the roots are mapped into the currently selected octave, but the tonic may not be not the lowest note in the range in all keys. It should be possible to arrange the 'fold point' of the octave arbitrarily, and experiment with other schemes that do not map all of the roots into a single octave.

Chord quality schemes for common modes not linked to window shapes. Only the major mode key window shape is currently inplemented. It would be useful to implement a harmonic minor window shape to go with and motivate the harmonic minor chord quality scheme, as described in Appendix 2.

---

[4] The reason for this will probably be familiar to anyone who has programmed graphics in a hurry. The note circles are illuminated and then erased by copying them once, and then twice, in standard graphics XOR transfer mode. The bug would be trivial to fix.

4 Simple ways of improving the Harmony Space prototype

Apart from addressing the trivial programming bugs and missing features documented above, there are a number of simple measures that could greatly improve the usefulness of Harmony Space. An illustration of the kind of way in which these measures could be addressed is given in Chapter 5 and appendix 2. A small selection of these measures, and one or two new ones, are noted below

Re-implement secondary control functions
The secondary control functions, e.g.control of chord-arity, quality overide etcetera should be re-implemented in a DM fashion with menus or buttons, visible state and so on.

Alternative displays to aid skill transfer. It would be trivially to arrange for all notes played to appear on a piano keyboard display as they sounded (see appendix 2). (This could even be done with commercial programs or hardware connected to Harmony Space via MIDI.) This measure might help with transfer of skills to keyboard instruments. Similarly, it would be trivially possible for all notes played to appear on a staff notation display (see appendix 2). This might be best displayed using notional, uniform note durations. (Actual durations could be used but might soon look very messy .)

Rubber banding or numbering for progression tracing. When reviewing the trace of a chord sequence, it can be hard to see the order of the chords. It would be trivial but useful to arrange for arrows to be shown on screen linking consecutive notes, or perhaps to number the chords on the display.

User defined chord qualities. It would be trivial but useful to arrange for users to be able to define their own chord qualities or chord quality schemes by example.

Better rhythmic control. It would be useful to allow rhythmic patterns read in from a MIDI keyboard or note editor to be combined with chord progressions when desired. Alternatively, it might be useful to have rhythmic figures selectable from a menu triggered at each button depression. Yet again, it might be useful to use a drum pad combined with Harmony Space like a version of the intelligent

drum (see chapter 4). Finally, a midi glove (see chapter 4) might be a better controller than a mouse, since finger gestures could be use to control rhythmic patterns for each chord or arpeggiation.

Switchable Balzano/Longuet-Higgins array. From an educational and psychological point of view, the two versions of Harmony space (based on the Balzano and Longuet-Higgins theories respectively) have interesting and subtlety different implications, as explored in Chapter 7. However, from an implementational point of view, the two versions of Harmony space are identical except for a shearing of the key-windows and a shearing of the layout of the note circles. It would be easy and useful to arrange for a single interface to support both layouts.

For other possible improvements to the Harmony Space prototype , see Chapter 5 and appendix 2.

5 Conclusions

The prototype implementation is rough and ready but permits simple experiments to be carried out. The prototype demonstrates the coherency and consistency of the key elements of the design of Harmony Space.

6 Summary of commands supported by Harmony Space prototype

The following table summarises all of the commands available in the prototype implementation of Harmony Space. The annotation 'toggle' indicates that the same key alternately switches a feature on and off.

Action		Function Key
Starting and finishing		
exit Harmony Space		Exit
redraw screen		Again
re-initialise midi interface		Read
Display management		
display whole chord or root only	(toggle)	Edit
leave trail or only show current event	(toggle)	Backspace
Key window		
move key window		up, down, left, right and diagonal arrow keys
Octave control		
octave up		Mark
octave down		Char Delete
set octave to middle C area		Line delete

Standard chord-arities

single notes	F1
diads (scale-tone fifth version)	F2
triads	F3
sevenths	F4
scale tone ninths (III and VII are played as sevenths)	F5

Common variant chord quality schemes for the major mode

diads (scale-tone thirds version)	<shift> F2
triads (version with dominant seventh on V)	<shift >F3

Common chord quality schemes for harmonic minor

triads (version with V of harmonic minor as major quality)	F7
sevenths (version with V of h. minor as dominant seventh)	<shift> F7
triads (version with V of h. minor as dominant seventh)	<ctrl> F7

Common chord quality schemes for dorian mode

triads (version with V of dorian as major quality)	F8
sevenths (version with V of dorian as dominant seventh)	<shift> F8
triads (version with V of dorian as dominant seventh)	<ctrl> F8

Overriding chord quality

major	1
dominant seventh	<shift> 1
minor	2
minor seventh	<shift > 2
diminished	3
diminished seventh	<shift> 3
dominant seventh #3	4
13th major minor third	5
major sixth	6

Switching automatic mechanism off and on

Freeze chord quality	(toggle)	Hold

Setting chord quality of chromatic roots

set quality of chromatic roots to last quality	Pop
mute chromatic roots	<shift> F1

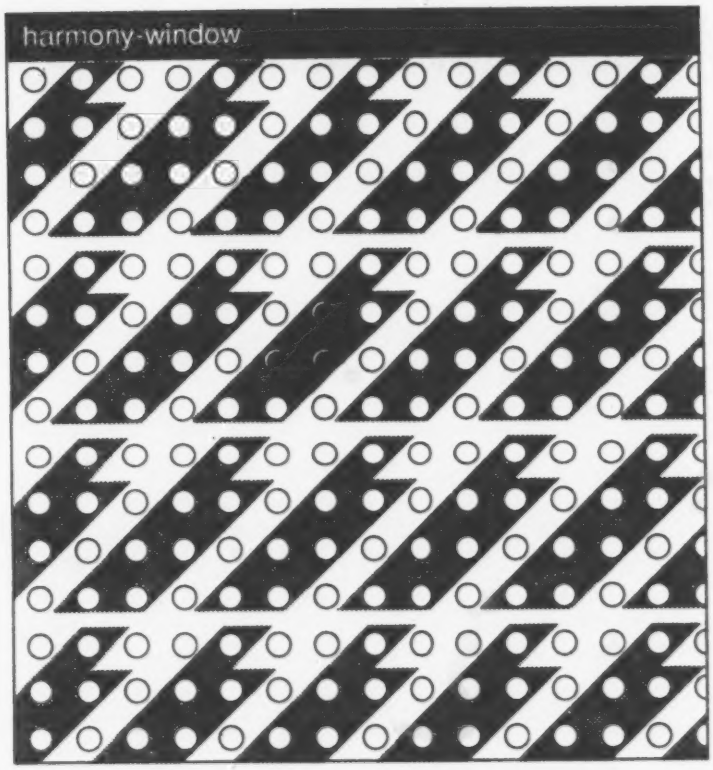


Fig 1

A screen dump of tonic major triad being sounded on the Harmony Space prototype[5].

[5] As noted earlier, for technical reasons to do with the the difficulty of printing large expanses of black, we have presented all of the screen dumps in this appendix inverted black-for white. Since the actual screen dumps are over 1Mb each in size - too large to include in the thesis document - they have been redrawn manually as accurately as possible.

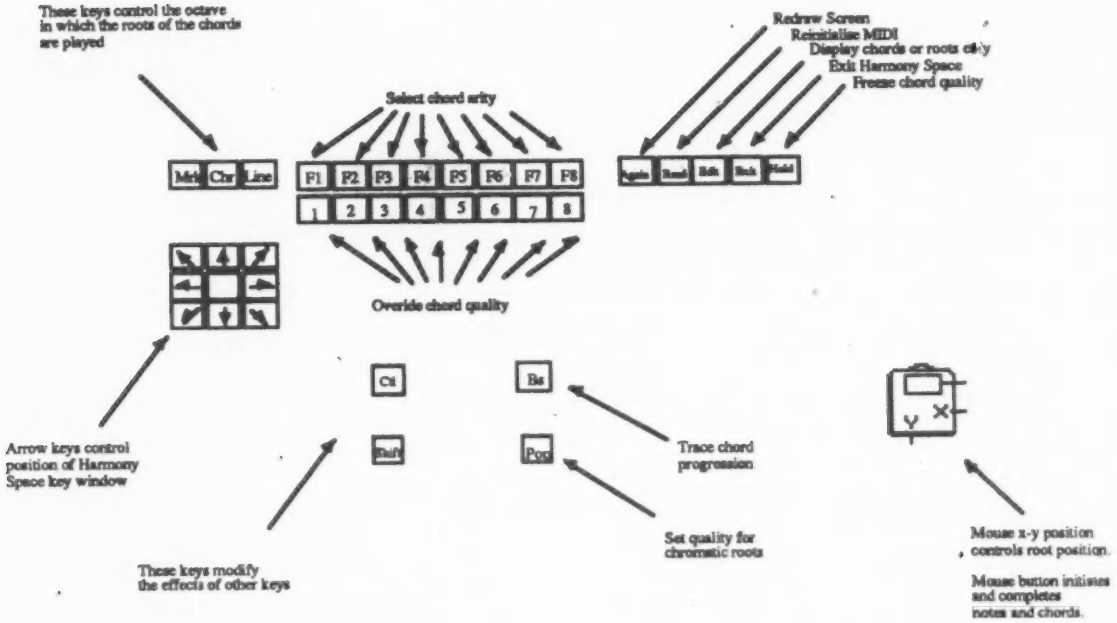


Fig 2

Overview of layout of function keys in prototype implementation of Harmony Space

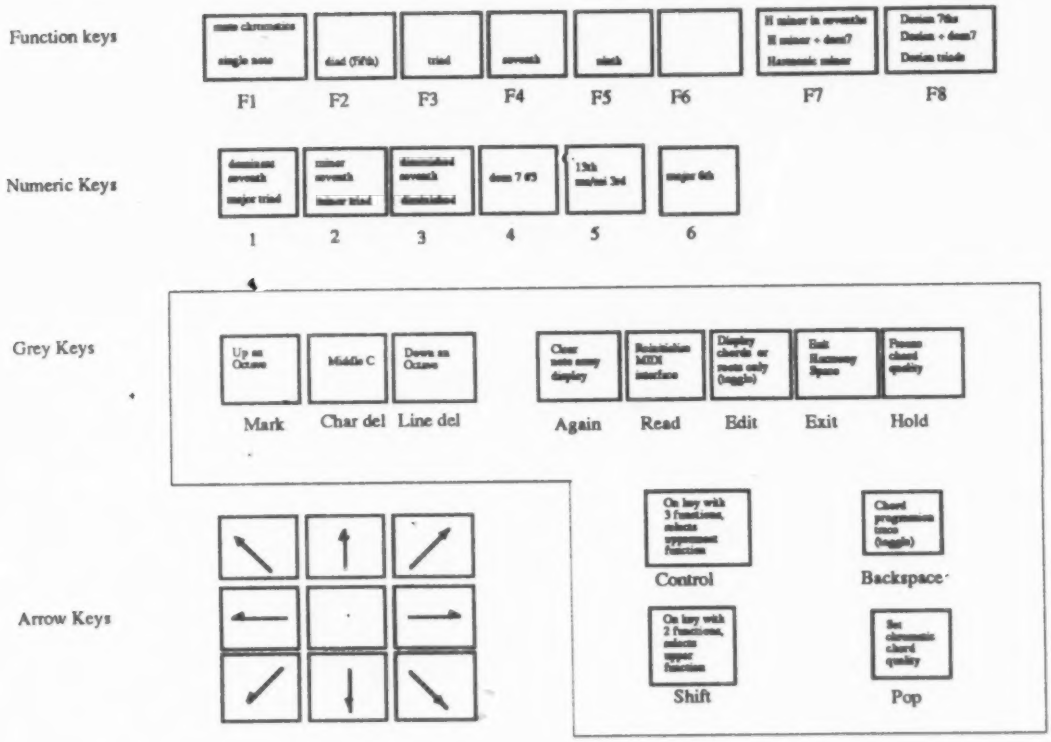


Fig 3

Function of each key in prototype implementation of Harmony Space

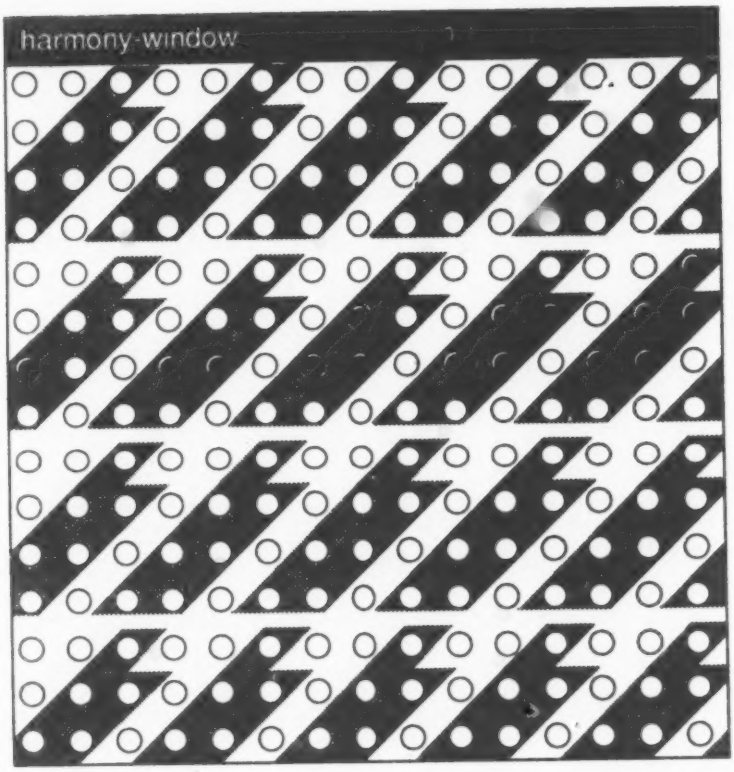


Fig 4

Screen dump from prototype illustrating five stages of scale-tone chord construction: note, diad,triad seventh and ninth..
Note that the option *not* to erase chord traces automatically after they have been sounded has been selected.

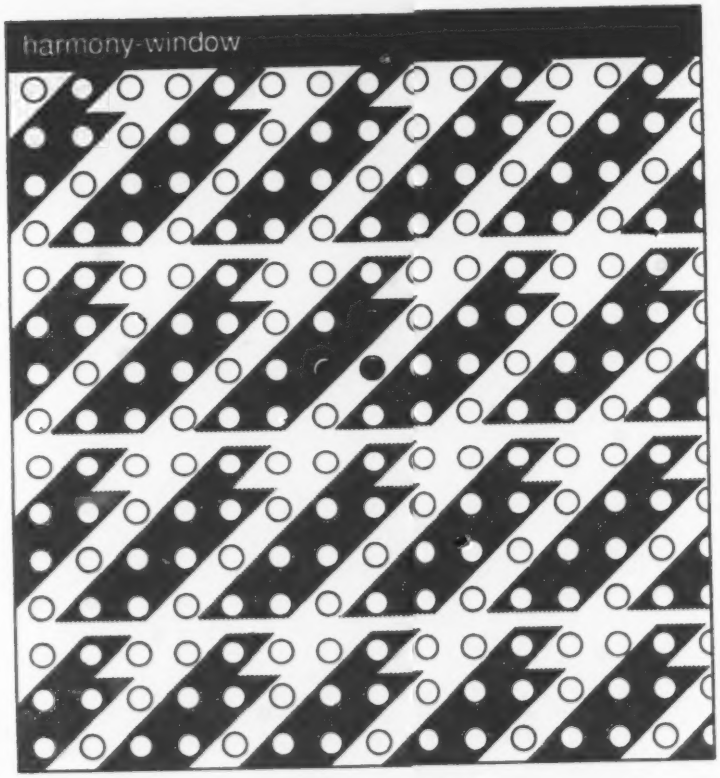


Fig 5

Screen dump from prototype illustrating the playing of a V chord in the harmonic minor mode. Note that the major key quality does *not* conform to the diatonic key-window shape.

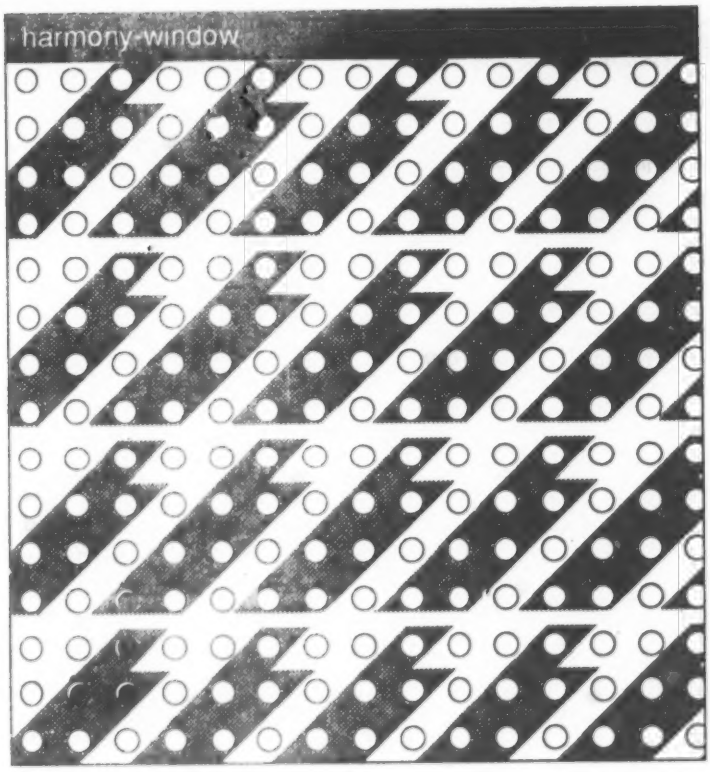


Fig 6

The first of three screen dumps from the prototype illustrating the automatic change of chord quality associated with a fixed note as the key window is moved.

The chord-arity used for the illustration is the scale-tone seventh chord-arity (i.e. four-note scale-tone chords).

In this figure, the note occupies the position in the key window corresponding to the degree of the scale V.

In order to fit the key-window in accordance with the chord-growing rule, the chord quality is dominant seventh.

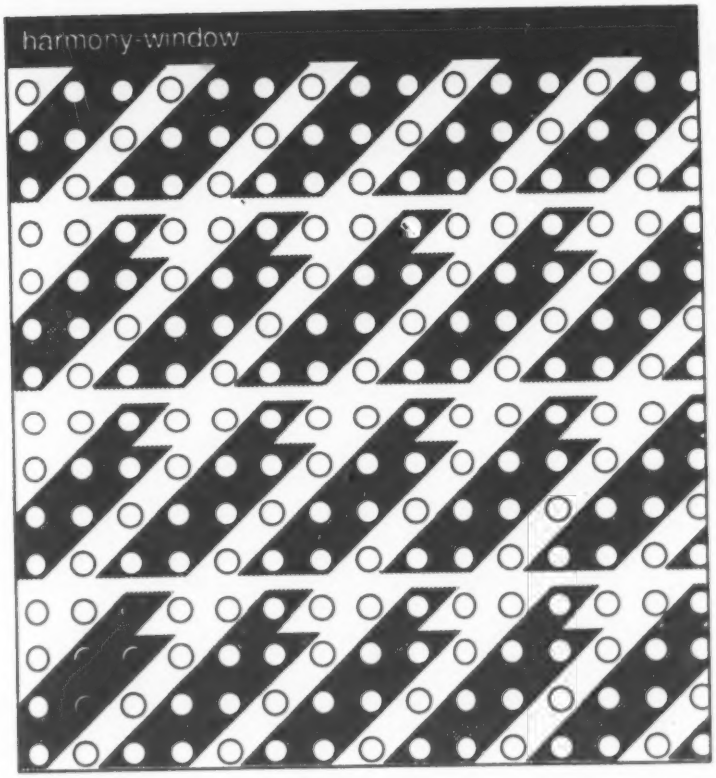


Fig 7

The second of three screen dumps from the prototype illustrating the automatic change of chord quality associated with a fixed note as the key window is moved.

In this figure, the key window has shifted vertically down a minor third relative to the original key.

The fixed note now occupies the position in the key window corresponding to the degree of the scale III in the new key.

In order to fit the key-window in accordance with the chord-growing rule, the chord quality is minor seventh.

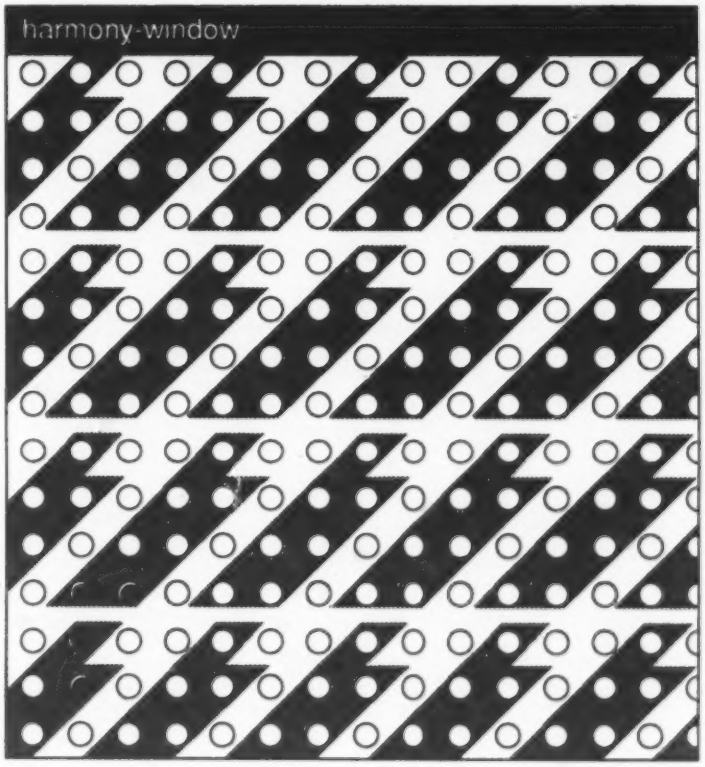


Fig 8

The final of three screen dumps from the prototype illustrating the automatic change of chord quality associated with a fixed note as the key window is moved.

In this figure, the key has shifted (horizontally) down a major third relative to the original key.

The note now occupies the position in the key window corresponding to the degree of the scale VII of the new key.

In order to fit the key-window in accordance with the chord-growing rule, the chord quality is diminished seventh.

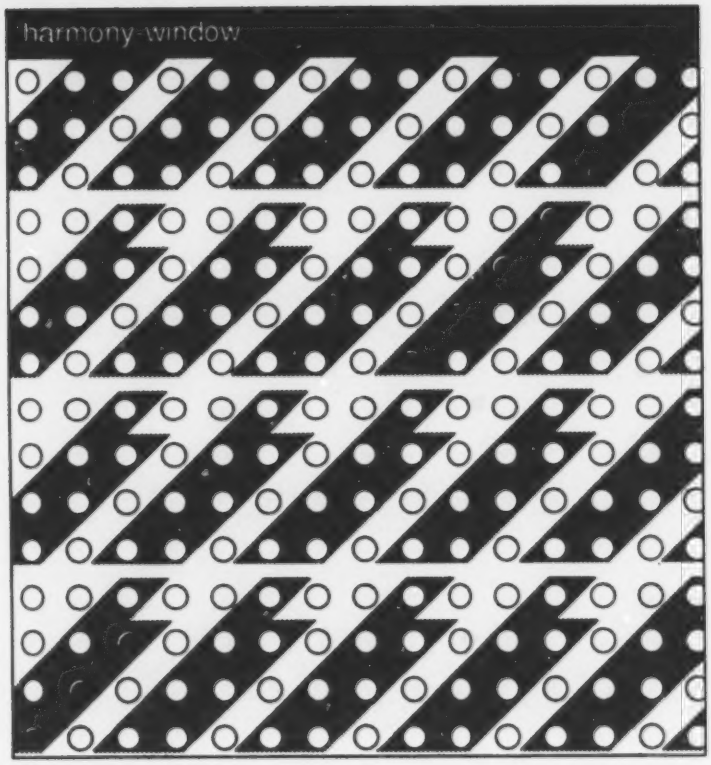


Fig 9

Screen dump from prototype illustrating the playing of the chord sequence of Street Life (Crawford, 1978) (see chapter 9 for chord sequence).

The option *not* to erase chord traces automatically after they have been sounded is in force.

The option to display the roots of chords only (as opposed to the whole chord) has also been selected

Appendix 2 : Harmony Space notional interface

We noted in chapter 5 and appendix 1 that the prototype implementation of Harmony Space lacks proper menus, labelling, and any almost other than the minimal features described in chapter 5[1]. In this appendix, we present a specification for a slightly less minimal version of Harmony Space, referred to as the 'idealised' version. The arrangements suggested in this appendix are not claimed to be ideal - they are offered as an illustration. This hypothetical version of Harmony Space requires little in the way of non-standard implementation techniques and covers little new conceptual ground, but would almost certainly be much easier to use. To avoid continual circumlocution, we will refer to the idealised version as though it exists, although it is currently only a specification. The idealised version conforms to the minimal specification of Harmony Space given in chapter 5, with a number of straightforward extra features added: to give one or two examples, there is an extra mouse and foot switches in place of function keys; the note circles are labelled; repositionable ('tear-off') menus are used to make the interface easier to use, and so on. The other major difference between the notional version and the implemented prototype is that it is reconfigurable to display a screen array conforming with either the Longuet-Higgins or the Balzano theory. Since appendix 1 uses the Balzano theory exclusively, for presentational balance we will give illustrations in this appendix in the Longuet-Higgins configuration. The notional interface is an illustration: no particular claims are made for the menu layout, etcetera. We begin with an illustrated overview of the appearance and method of use of the notional interface.

1 General appearance and operation of prototype

The Harmony Space window for the idealised version is shown in Fig 1. In the

[1] Note - the material in this appendix inevitably overlaps to some extent material already seen in Chapter 5 and appendix 1. Identical or almost identical wording is used in parts, although the description is carefully tailored here to refer to the notional interface as opposed to the prototype or minimal version.

Longuet-Higgins configuration, notes are arranged in ascending major thirds (semitone steps of size 4) on the x-axis, and in ascending perfect fifths (semitone steps of size 7) on the y-axis. Unlike in the prototype, the user can direct for the note circles to be labelled either in terms of alphabetic note names (fig 1), roman numerals, (fig 5) or semitone numbers (not illustrated) by means of selections on the graphic display menu (fig 3). The key windows can be slid as a group over the fixed grid of circles by means of a mouse, midi-glove, People Mat[2] or other pointing device (see fig 2 for a set of suggested controllers and their suggested functions). A different pointing device (we will just say 'mouse' in future to avoid clumsy phrasing) is used to move the cursor around in the window.

(In this paragraph, we note features of the mouse action that apply in all versions of Harmony Space.) If the mouse button is pressed while the cursor is over a note circle, the circle lights up and the corresponding note sounds audibly until the button is released. If the button is *kept* depressed and the mouse is swept around, the appropriate succession of new notes is sounded and completed as the mouse enters and leaves the note circles. More generally, the mouse can be set to control the location of the root of a diad, triad, seventh or ninth chord. For any given chord arity, depressing the mouse button on a root causes all of the notes of the appropriate scale-tone chord to light up and sound. (Fig 1 shows a C major triad being sounded in this way).

Apart from the two mice, other functions in the idealised interface are controlled by a set of foot switches, and a set of menus. (See figs 2, 3 and 4 for more details. Fig 2 is an overview of suggested gestural controllers and their suggested functions, Fig 3 shows the major menus, and fig 4 shows the secondary menus. There is also a summary discussion of menu functions at the end of this appendix.) The sliding foot switch (see fig 2) allows the user to select whether she wants at any time chords of arity one, two, three, four, five, etcetera. (i.e., single notes, diads, triads, sevenths, ninths). This foot switch is designed to allow the user to play interspersed single notes and chords fluidly. As in all versions of Harmony Space, the quality of each chord is chosen automatically (unless the user directs otherwise) in accordance with the prevailing chord-arity and the current location of the root relative to the current

[2] People Mats™ are commercially available RS232 pointing devices controlled by walking on them (Elsom-Cook,1986). See chapter 4 and appendix 1 for more notes on midi gloves.

position of the key window. This automatic assignment of quality to chords can be overridden at any time by selecting the desired quality using the 'quality overide' menu (fig 3). The second pointing device (used mainly for controlling modulation) is also used to control the overiding of chord quality.

Many other controller arrangements are possible. For example, a foot switch could be used to choose between a limited number of chord qualities pre-selected by the user. Alternatively, control of modulation could be assigned to a mole (i.e. a foot-controlled two-dimensional pointing device (Pearson and Weiser, 1986)), leaving the second hand free to deal with chord quality interventions. Ideally, the user should be allowed to patch arbitrary control devices to control arbitrary parameters (e.g chord inversion and key-window shape), along the lines of the 'M' program reviewed in Chapter 4.

The remainder of this appendix is taken up with a summary discussion of each menu function, grouped by menu.

2 Summary of Menu functions

2.1 Conventions common to all menus

The user can choose how particular menus are to behave. The menus can behave either in conventional 'pull-down' mode or can stay 'pulled down' until further notice for easier access. This feature is toggled by double clicking on the menu bar. This allows the user to customise the interface in accordance with which menus are to be used most intensively. (Menus perform here, as elsewhere, a triple reduction of cognitive load: they show what choices are available; show the current choices in force; and allow choices not under consideration to be hidden.) In figs 2 and 3, following the Macintosh menu convention, dots (...) after a menu entry indicate that if the entry is selected, a secondary dialogue will follow to elicit additional information.

2.2 Graphic display menu

The *show root only* command, and its inverse *show whole chord* control whether the entire chord or just its root is displayed (as exemplified by figs 1 and 6 respectively). The audible result remains unaffected either way.

The command *show trail* and its inverse *no trail* control whether the trail of notes that have been played are erased as the chords are released (fig 1) or left on the screen to show the trail of the progression (fig 6). Experiments with the prototype have shown that trails can get very messy with the 'show whole chord' option on: trails are generally more useful in combination with the 'show root only' command. To rubout all trails, the command *tidy screen* is used.

The *show keyboard* command and its inverse *hide keyboard* control whether or not the notes being played at any moment are shown on a keyboard display (see fig 1). This could be simulated with the current prototype of Harmony Space by connecting it via midi to commercially available equipment (personal communication, Desain and Honing).

The *show staff* command and its inverse *hide staff* control whether or not the notes being played are cumulatively displayed on a staff display (fig 1). This can be simulated on the current prototype of Harmony Space by connecting it via midi to commercially available programs such as 'Personal Performer'. In practice, the varying durations of the chords can make the display rather messy. The command *hide staff duration* causes all of the durations to be shown uniformly (fig 5). To clear the staff, the command *tidy staff* is used.

The history of the chord sequence may be displayed beneath the harmony window as a trace of alphabetic chord names (figs 5 and 6) or roman numeral chord names (not shown) using the commands *show alpha history* and *show roman history* . In either case, modulations are indicated both textually and using typographic shifts (fig 5). To clear the history, the commands *tidy roman history* and *tidy alpha history* are used.(The prototype Harmony Logo interpreter maintains multiple views of chord sequences of a similar sort which can be printed out to help the user analyse chord sequences. The ease of providing and modifying such views was one of the

major motivations for using an object-oriented style to implement Harmony Logo.)

The *show key name* command simply causes the key that Harmony Space is currently in to be displayed textually to complement the graphic display (not illustrated).

The commands *alphabetic labels, roman numeral labels* and *semitone number labels* are used to control the labelling of the note circles (see figs 1 and 5).

Finally, the *highlight centre* command controls whether or not the label of the currently selected harmonic centre (see harmonic centre menu) is emboldened in the note-circle labels, as it is in Fig 5 (although this has not reproduced strongly on the page).

2.3 Overide quality menu

Qualities selected on this menu override the automatically selected chord quality (just for the next event to be triggered). When the *quality lock* command is selected, the quality of the chord is locked to its previous value (or a value selected on the menu just before the quality lock) until the quality lock is switched off. The *define new chord quality* option allows a user to define a new chord quality by example on the screen and name it, after which its name will appear on the quality overide menu like that of any other quality. The *remove new chord quality* option allows such entries to be removed.

2.4 Chord-arity menu

The *chord-arity* menu decides the size of the chord (one, two, three, four , five notes) triggered at any time by a single mouse button depression. The quality of the resulting chord depends not only on the mouse location (i.e. the degree of scale of the root) and the key window location (i.e. the key), but also on the chord quality scheme in force (see entry for chord quality scheme menu).

2.5 Filter and alphabet menus

When the key filter is *on*, the mouse pointer is constrained to move exactly

vertically, horizontally or diagonally at all times; movement at intermediate angles is interpreted as whichever of the constrained directions it most closely fits. With the key filter on, when a mouse movement would take the cursor to a note circle outside the key window (i.e to a note not in the current key), the cursor jumps to the next square on the same bearing which is in a key window. To reflect the state of the key filter graphically to the user, when the key filter is off the outline of the key window is drawn in a dotted line, as opposed to the continuous line used when the filter is on. The option *tritone wrap* simply means that transitions from IV to VII are displayed in a single key window (as in the traces shown in Chapter 5, fig 6). Note that this option does not prevent chromatic notes from being sounded as part of a chord - it only prevents them from acting as roots. The *alphabet filter* works in exactly the same way as the key filter, but instead of allowing only roots in the key to be played, it allows only roots selected in the alphabet menu to be played. The option *window wrap* means that all chords are displayed in a single representative of the key window.

2.6 Window shape menu

The window shape menu allows key windows of different shapes appropriate to the harmonic minor and other modes to be displayed. Selecting a window shape automatically causes the corresponding harmonic centre to be selected in the harmonic centre menu, and the corresponding chord quality scheme to be selected in the chord quality menu, although this may be overridden using the latter two menus independently. New window shapes can be specified by the user, and associated with particular harmonic centres and chord quality schemes using the *define new shape* option.

2.7 Harmonic centre menu

The choice of harmonic centre (I to VII) corresponds to specifying that one of the following modes: ionian (or major), dorian, phrygian, lydian, mixolydian, aeolian (or minor) or locrian is in use. This makes no difference to the chords audibly produced, (except in two circumstances: if the chord quality scheme is also changed - see the entries for the 'window shape' menu and 'chord quality scheme' menu - or if a modal pedal is selected - see the entry for the 'extras' menu). Choice of harmonic

centre also affects the harmony window display in the following ways. Firstly, it affects which root name is emboldened when the 'highlight centre' command is used on the graphic display menu (see the entry for the graphic display menu). Secondly, it affects the assignment of roman numerals to note names when labelling the screen with roman numerals. Thirdly, it affects the roman numeral history (see the notes on 'roman numeral history' under the entry above for the graphic display menu).

2.8 Chromatic quality menu

The quality of chords associated with chromatic notes in their capacity as potential roots of chords may be fixed using the 'chromatic quality' menu. The same quality applies to all chromatic notes. As a special case, chromatic notes may be muted (this does not affect their sounding as non-root notes in chords where appropriate.

2.9 Chord quality scheme menu

The chord quality scheme menu allows different degrees of the scale to be set to trigger chord qualities other than the scale tone qualities. Note that changes in the window shape automatically make changes in the chord quality scheme selected, but these may be overridden. The user may define and remove new chord quality schemes using the *define new scheme* and *remove scheme* options.

2.10 Midi menu

The midi menu simply allows different synthesiser patches (timbres) to be selected.

2.11 Extras menu

If the option *doubled root* is selected, each chord is accompanied by a note in the bass doubling the root. If the option *modal pedal* is selected, each chord is accompanied by a modal pedal note appropriate to the harmonic centre in force. The option *octave range* allows an upper and lower octave range to be assigned to the pitch of chord roots. Limits can be assigned separately for the x and y axes. So for example, if the x-axis is assigned a lower limit of octave 4 and an upper limit of octave 6, moving the cursor along the x-axis gradually steps the root through two octaves as a natural side-effect of adding major thirds. On coming to the top of the second octave, the pitch 'wraps around' to the bottom octave. This is repeated along

the x-axis. The option *bass range* is identical to the option 'octave range' except that whereas 'octave range' applies to the root, 'bass range' applies to the doubled root or modal pedal. The options *root position, first inversion, second inversion* and *third inversion* control the inversion of the chord sounded (leaving out of consideration any doubled root). The octave height of the notes in the chord in any inversion is determined by the octave height of the root. The *spacing* option allows any note in the chord to be exported up or down a specified number of octaves.

2.12 Rhythmic figure and bass figure

The rhythmic figure and bass figure menus allow a single mouse button depression to trigger rhythmic figures in the chord and doubled root or modal pedal independently.

2.13 Arpeggiate and alberti pattern menus

The *arpeggiate* menu allows a single mouse button depression to trigger one or more upward or downward arpeggios. These can be combined with rhythmic figures (see the notes on the 'rhythmic figure' menu). The *alberti pattern* option allows a single mouse button depression to trigger an arbitrary alberti figuration pattern. This option may be combined with the rhythmic figure option.

This completes the summary discussion of each menu function, and the discussion of the idealised version of Harmony Space. Diagrams for this document are appended at the end.

3 Conclusion

Despite the term 'idealised' interface, there is no suggestion that the arrangements suggested in this appendix are particularly well-thought out or ideal. The arrangements would need to be tested and improved by a process of trial and error. The purpose of this appendix has been to illustrate one way in which a fuller implementation of Harmony Space might be designed.

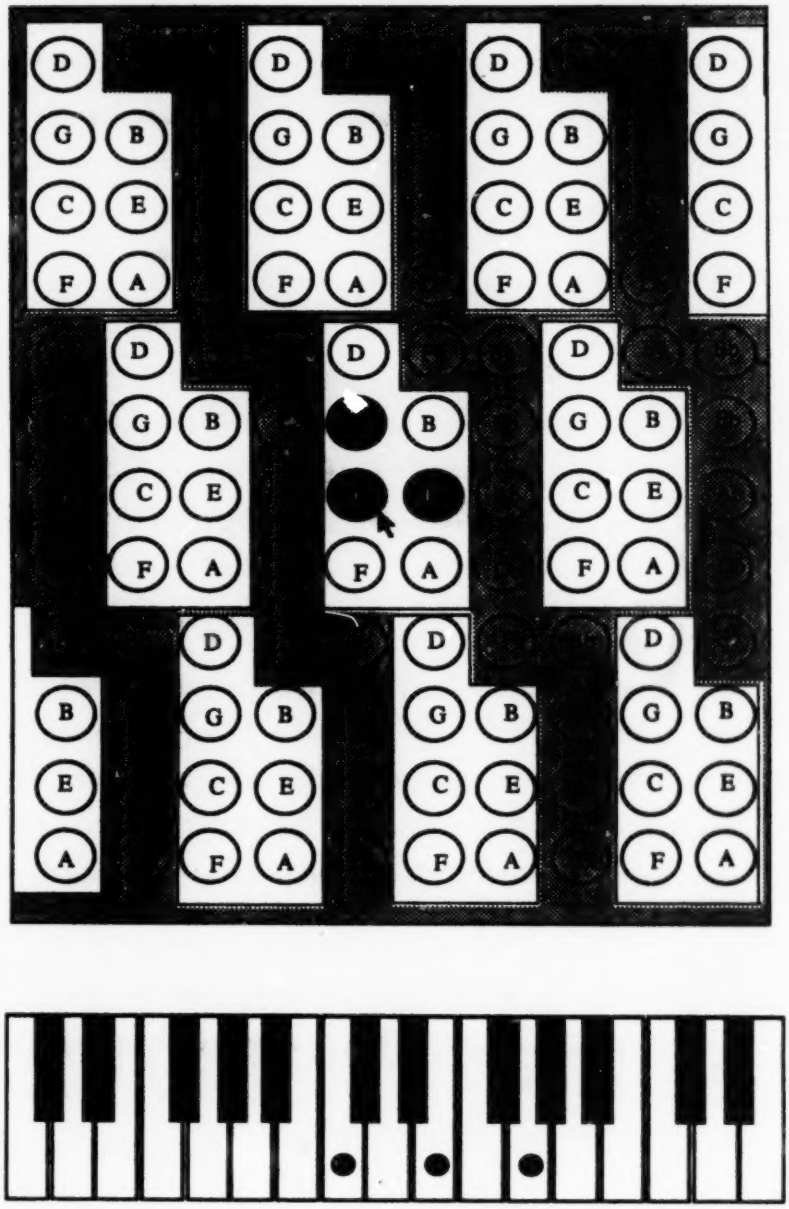
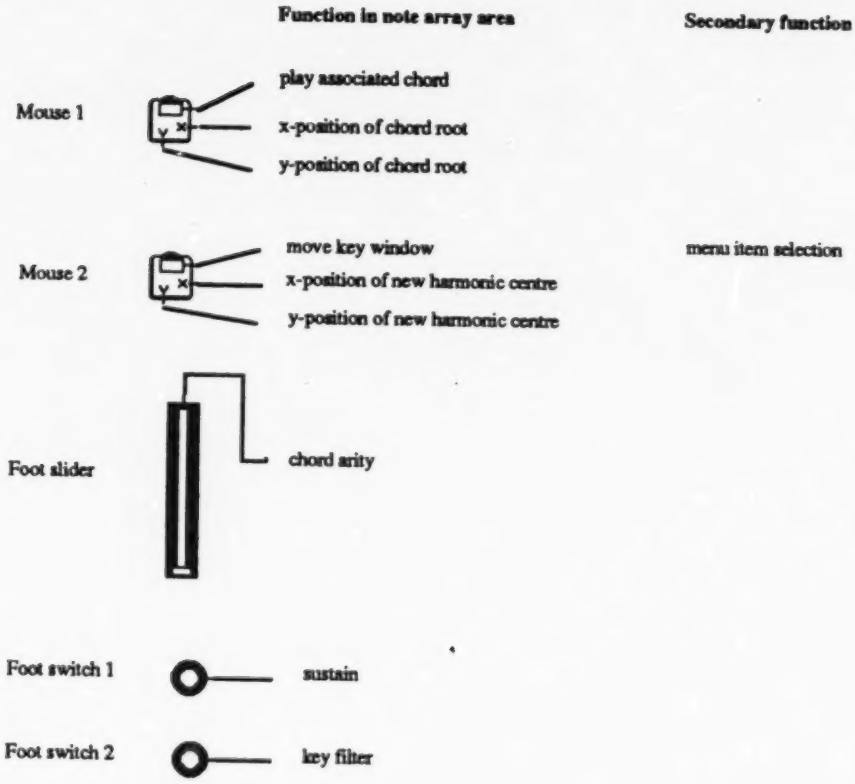


Fig 1
Idealised version of Harmony Space with keyboard display.



Summary of suggested controller functions

Mouse 1	play notes
Mouse 2	move key window & overide chord quality
Foot slider	change chord arity
Foot switch 1	key filter on/off
Foot switch 2	sustain

Fig 2
Suggested gestural controllers and their functions in idealised version of Harmony Space.

Fig 3

Most commonly used menus for idealised version of Harmony Space.

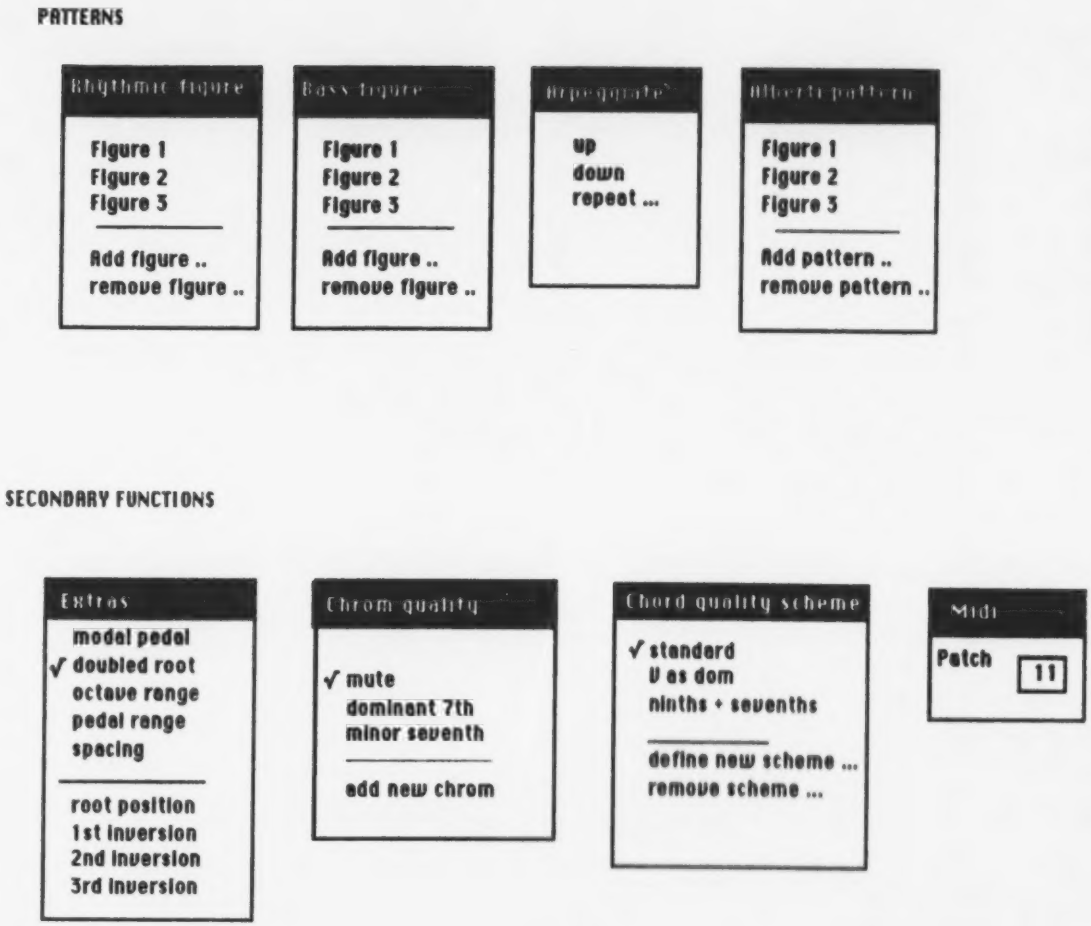
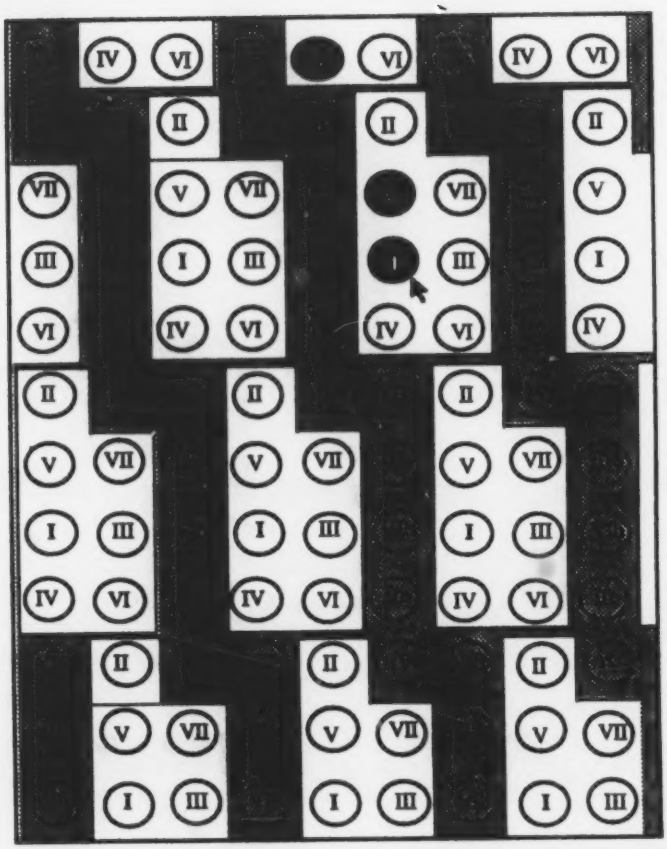


Fig 4
Secondary menus in idealised version of Harmony Space.



C major: A mi7 D mi7 G dom7 C ma7 F ma7
E major: B dom7 Ema7

Fig 5

Display from idealised version of Harmony Space showing trace of first eight bars of "All the things you are" in C major, with only roots shown in the trace, roman numeral labels displayed, and an alphabet name history given.

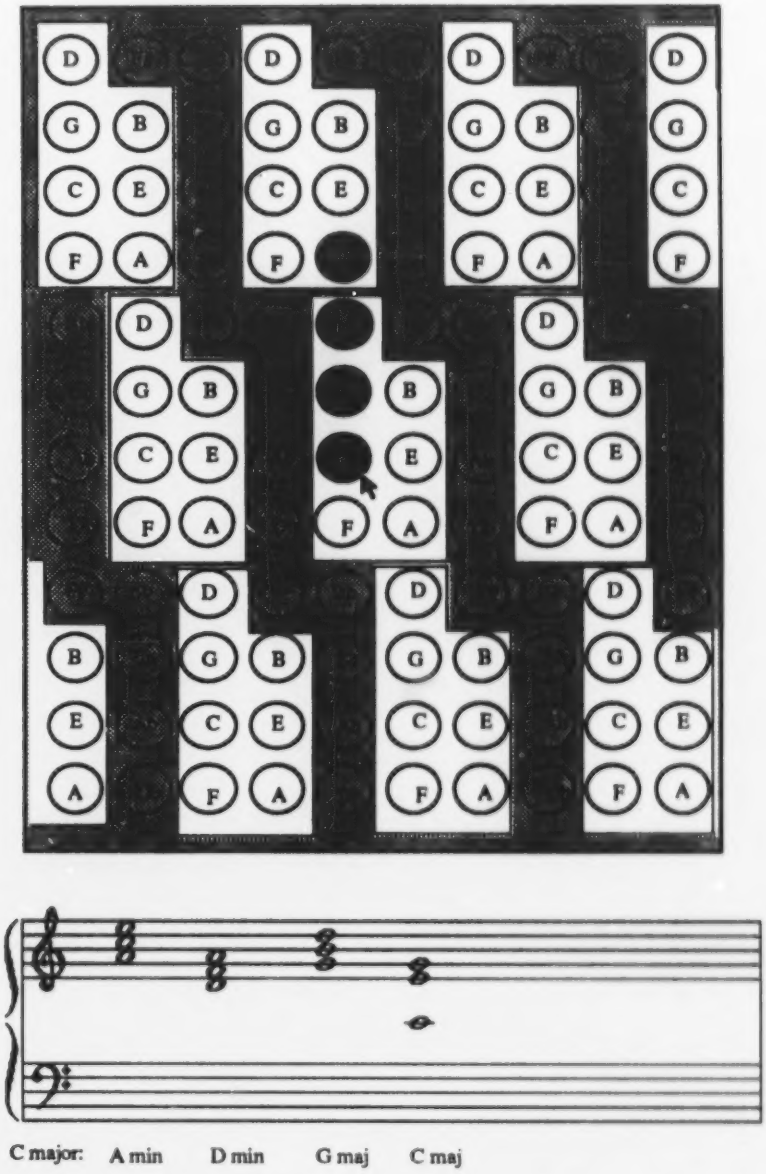


Fig 6
Display from idealised version of Harmony Space showing trace of "I got rhythm" progression with roots only displayed and staff notation display given.

Appendix 3: Chord symbol conventions.

Two chord notation conventions are used in the thesis, Mehegan's (1959) convention and the 'classical' convention. The main difference between the two conventions is that in the classical version, chord qualities are always given explicitly: in particular, upper case roman numerals are major, and lower case numerals are minor. Other aspects of chord quality are shown explicitly by annotation. In the Mehegan convention, all numerals are given in upper case, since the chord quality is understood to be the scale-tone chord quality appropriate to the degree of the scale of the key in force - only altered chords or chords with added notes need to have their quality indicated explicitly by annotation. We have trivially extended Mehegan's convention into the modal case so that in the Dorian mode, for example, D is notated as I. To distinguish between the two systems typographically, chord sequences notated using the classical convention are <u>underlined</u>, those using Mehegan's system are not.

More details on Mehegan's convention

Roman numerals representing scale tone triads or sevenths are written in capitals, **irrespective** of major or minor quality (e.g. I II II IV V etc.). Roman numerals represent triads of the quality normally associated with the degree of the tonality (or modality) prevailing. We call this quality the "default" quality. *In jazz examples, Roman numerals indicate scale-tone sevenths rather than triads.* The following post-fix symbols are used to annotate Roman chord symbols to override the chord quality as follows ; x - dominant, o - diminished, ø - half diminished, m - minor, M - major. The following post-fix convention is used to alter indicated degrees of the scale; "#3" means default chord quality but with sharpened 3rd; "#7" means default chord quality but with sharpened 7th, etc. The following post-fix convention is used to add notes to chords; e.g. "+6" means default chord quality with added

scale-tone sixth 6th. The prefixes # and b move **all** notes of the otherwise indicated chord a semitone up or down.

General points

In the thesis, inversions are mostly ignored, as discussed in chapters 5 and 9. Mehegan's convention that unannotated Roman numeral refer (by default) to scale tone sevenths has been extended.We allow unannotated Roman numerals to stand for the scale tone chord of any specified chord-arity (see Chapter 5 for an explanation of the Harmony Space concept of chord-arity).